

## Promotoren zur Expression von Genen in Tagetes

## Beschreibung

5 Die vorliegende Erfindung betrifft die Verwendung von Promotoren zur Expression, vorzugsweise zur blütenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes, die genetisch veränderten Pflanzen der Gattung Tagetes sowie ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung der genetisch veränderten Pflanzen.

10 Verschiedene biosynthetische Produkte, wie beispielsweise Feinchemikalien, wie unter anderem Aminosäuren, Vitamine, Carotinoide, aber auch Proteine werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik-, Feed-, Food- und pharmazeutischen Industrie.

Diese Substanzen, die zusammen als Feinchemikalien/Proteine bezeichnet werden, umfassen unter anderem organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Nukleotide und Nukleoside, Lipide und Fettsäuren, Diöle, 20 Kohlenhydrate, aromatische Verbindungen, Vitamine, Carotinoide und Cofaktoren, sowie Proteine und Enzyme. Ihre Produktion im Großmaßstab erfolgt zum Teil mittels biotechnologischer Verfahren unter Verwendung von Mikroorganismen, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren.

25 Carotinoide werden de novo in Bakterien, Algen, Pilzen und Pflanzen synthetisiert. In den letzten Jahren wird zunehmend versucht, auch Pflanzen als Produktionsorganismen für Feinchemikalien, insbesondere für Vitamine und Carotinoide zu nutzen.

30 Ein natürliches Gemisch aus den Carotinoiden Lutein und Zeaxanthin wird beispielsweise aus den Blüten von Marigold Pflanzen (Tagetes Pflanzen) als sogenanntes Oleoresin extrahiert. Dieses Oleoresin findet Anwendung sowohl als Inhaltsstoff von Nahrungsergänzungsmitteln als auch im Feed-Bereich.

35 Lycopin aus Tomaten findet ebenso Anwendung als Nahrungsergänzungsmittel, während Phytoen überwiegend im kosmetischen Bereich verwendet wird.

Ketocarotinoide, also Carotinoide, die mindestens eine Keto-Gruppe enthalten, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'- 40 Hydroxyechinenon, Adonirubin und Adonixanthin sind natürliche Antioxidantien und

Pigmente, die von einigen Algen, Pflanzen und Mikroorganismen als Sekundärmetabolite produziert werden.

5 Aufgrund ihrer farbgebenden Eigenschaften werden die Ketocarotinoide und insbesondere Astaxanthin als Pigmentierhilfsstoffe in der Tierernährung, insbesondere in der Forellen-, Lachs- und Shrimpszucht verwendet.

10 Ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von natürlichen, biosynthetischen Produkten und insbesondere Carotinoiden ist daher von großer Bedeutung.

WO 0032788 beschreibt einige Carotinoid Biosynthesegene aus Pflanzen der Gattung Tagetes und offenbart, wie genetisch veränderte Pflanzen der Gattung Tagetes hergestellt werden könnten, um in den Petalen verschiedene Carotinoidprofile zu erhalten  
15 und damit gezielt bestimmte Carotinoide herzustellen. Dazu ist es nötig, einige Biosynthesegene überzuexprimieren und andere zu unterdrücken.

Zur Überexpression der neu gefundenen Carotinoid-Biosynthesegene in Pflanzen der Gattung Tagetes wird in WO 0032788 der petalenspezifische Promotor der Ketolase  
20 aus Adonis vernalis postuliert.

Aufgrund einer Vielzahl möglicher Schwierigkeiten bei der Überexpression bestimmter Gene besteht ein ständiges Bedürfnis, weitere Promotoren zur Verfügung zu stellen, die eine Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.  
25

Der Erfindung lag daher die Aufgabe zu Grunde, weitere Promotoren zur Verfügung zu stellen, die die Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.

Demgemäß wurde gefunden, dass sich die Promotoren, ausgewählt aus der Gruppe  
30

- A) EPSPS Promotor
- B) B-Gene Promotor
- C) PDS Promotor und
- D) CHRC Promotor

35 sehr gut zur Expression von Genen in Pflanzen der Gattung Tagetes eignen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

Die Erfindung betrifft daher die Verwendung eines Promotors, ausgewählt aus der Gruppe

- A) EPSPS Promotor
- 5 B) B-Gene Promotor
- C) PDS Promotor und
- D) CHRC Promotor

- zur Expression von Genen in Pflanzen der Gattung Tagetes, mit der Maßgabe, dass
- 10 Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

- Benfey et al. (Plant Cell Volume 2, pp. 849-856) beschreiben den EPSPS Promotor aus Petunia als petalenspezifischen Promotor zur Expression von Genen in Petunia
- 15 hybrida.

- Ronen et al. (PNAS Volume 97, Number 20, 11102-11107 beschreiben den B-GENE Promotor aus Tomate als blütenspezifischen Promotor zur Expression von Genen in Tomaten.
- 20

- Corona et al. (Plant Journal Volume 9 Number 4 pp. 505-512), Mann et al. (Nature Biotechnology Volume 18 pp. 888-892) und Rosati et al. (Plant Journal Volume 24 Number 3 413-419) beschreiben den PDS Promotor aus Tomate als frucht- und blütenspezifischen Promotor zur Expression von Genen in Tomaten und Tabak.
- 25

- Vishnevetsky et al. (Plant Journal Volume 20 Number 4 pp. 423-431) beschreiben den CHRC Promotor aus Gurke als blütenspezifischen Promotor zur Expression von Genen in Gurke, und weiteren Pflanzen wie z.B. Nelke, Sonnenblume, Tabak.

- 30 Es sind weiterhin zahlreiche blütenspezifische Promotoren aus verschiedenen Organismen in der Literatur bekannt. Dabei wurde überraschend festgestellt, dass viele dieser Promotoren in Pflanzen der Gattung Tagetes nicht zur Expression, insbesondere nicht zur blütenspezifischen oder petalenspezifischen Expression von Genen führen.

- 35 Es war daher überraschend, dass sich die Promotoren, ausgewählt aus der Gruppe

- A) EPSPS Promotor
- B) B-Gene Promotor

C) PDS Promotor und

D) CHRC Promotor

- 5 sehr gut zur Expression, insbesondere zur blütenspezifischen und besonderes bevorzugt zur petalenspezifischen Expression von Genen in Pflanzen der Gattung *Tagetes* eignen.

- 10 Unter einem Promotor wird erfindungsgemäß eine Nukleinsäure mit Expressionsaktivität verstanden, also eine Nukleinsäure verstanden, die in funktioneller Verknüpfung mit einer zu exprimierenden Nukleinsäure, im folgenden auch Gen bezeichnet, die Expression, also die Transkription und die Translation dieser Nukleinsäure oder dieses Gens reguliert.

- 15 Unter „Transkription“ wird erfindungsgemäß der Prozess verstanden, durch den ausgehend von einer DNA-Matrize ein komplementäres RNA-Molekül hergestellt wird. An diesem Prozess sind Proteine wie die RNA-Polymerase, sogenannte Sigma-Faktoren und transkriptionelle Regulatorproteine beteiligt. Die synthetisierte RNA dient dann als Matrize im Prozess der Translation, der dann zum biosynthetisch aktiven Protein führt.

- 20 Unter einer „funktionellen Verknüpfung“ versteht man in diesem Zusammenhang beispielsweise die sequentielle Anordnung einer der erfindungsgemäßen Promotoren und einer zu exprimierenden Nukleinsäuresequenz und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der Nukleinsäuresequenz erfüllen kann. Dazu ist nicht  
25 unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu exprimierende Nukleinsäuresequenz oder das zu exprimierende Gen hinter (d.h. am 3'-Ende) der erfindungsgemäßen Promotorsequenz positioniert wird, so dass beide Sequenzen kova-  
30 lent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

- 35 Unter „Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate, verstanden.



Unter „spezifischer Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

5 Bei einer „verursachten Expressionsaktivität“ oder „verursachten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp so nicht vorhanden war.

10 Bei einer „erhöhten Expressionsaktivität“ oder „erhöhten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit die gebildete Menge des Proteins erhöht.

15 Die Bildungsrate, mit der ein biosynthetisch aktives Protein hergestellt wird, ist ein Produkt aus der Rate der Transkription und der Translation. Beide Raten können erfindungsgemäß beeinflusst werden und damit die Rate der Bildung von Produkten in einem Mikroorganismus beeinflussen.

20 Die Bezeichnung „dass Gene aus Pflanzen der Gattung *Tagetes*, die in Wildtyppflanzen der Gattung *Tagetes* von dem „jeweiligen“ Promotor exprimiert werden, ausgenommen sind“, bedeutet, dass beispielsweise der EPSPS Promotor aus Pflanzen der Gattung *Tagetes* nicht zur Expression von EPSPS-Genen aus Pflanzen der Gattung *Tagetes* verwendet wird. Dahingegen kann das EPSPS-Gen aus Pflanzen der Gattung *Tagetes* erfindungsgemäß durch einen B-Gen Promotor, PDS Promotor oder CHRC Promotor aus Pflanzen der Gattung *Tagetes* exprimiert werden.

25 Unter dem Begriff „Wildtyp“ oder „Wildtyppflanze“ wird erfindungsgemäß die entsprechende Ausgangspflanze der Gattung *Tagetes* verstanden.

30 Je nach Zusammenhang kann unter dem Begriff „Pflanze“ die Ausgangspflanze (Wildtyp) oder eine erfindungsgemäße, genetisch veränderte Pflanze der Gattung *Tagetes* oder beides verstanden werden.

35 Vorzugsweise wird unter „Wildtyp“ für die Erhöhung oder Verursachung der Expressionsaktivität oder Expressionsrate und für die Erhöhung des Gehalts an biosynthetischen Produkten die Pflanze *Tagetes erecta*, insbesondere die Pflanze *Tagetes erecta Hybrid 50011* (WO 02012438) als Referenzorganismus verstanden.

40 Unter einem „EPSPS Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine 5-Enolpyruvylshikimat-3-phosphatsynthase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder

durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

- 5 Diese EPSPS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen EPSPS Promotorsequenzen oder den Nukleinsäuren, kodierend eine 5-  
10 Enolpyruvylshikimat-3-phosphatsynthase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine 5-Enolpyruvylshikimat-3-phosphatsynthase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

- 15 Unter einer 5-Enolpyruvylshikimat-3-phosphatsynthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Shikimat-3-Phosphat in 5-Enolpyruvylshikimat-3-Phosphat umzuwandeln.

- 20 Bevorzugte EPSPS Promotoren enthalten

A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder

- 25 A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder 3 aufweist oder

A3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert oder

- 30 A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3)

- Die Nukleinsäuresequenz SEQ. ID. NO. 1 stellt eine Promotorsequenz der 5-Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus *Petunia hybrida* (AAH19653) dar.

- 35 Die Nukleinsäuresequenz SEQ. ID. NO. 2 stellt eine Promotorsequenz der 5-Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus *Petunia hybrida* (M37029) dar.

- Die Nukleinsäuresequenz SEQ. ID. NO. 3 stellt eine weitere Promotorsequenz der 5-Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus *Petunia hybrida* dar.
- 40

Die Erfindung betrifft weiterhin EPSPS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 1, 2 oder 3) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1,2 oder 3 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße EPSPS Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 1, 2 oder 3 leicht auffinden.

Künstliche erfindungsgemäße EPSPS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

Die folgenden Definition und Bedingungen der Identitätsvergleiche und Hybridisierungsbedingungen gelten für alle Nukleinsäuren, also alle Promotoren und Gene der Beschreibung.

Unter dem Begriff "Substitution" ist der Austausch einer oder mehrerer Nukleotide durch ein oder mehrere Nukleotide zu verstehen. „Deletion“ ist das Ersetzen eines Nukleotides durch eine direkte Bindung. Insertionen sind Einfügungen von Nukleotiden in die Nukleinsäuresequenz, wobei formal eine direkte Bindung durch ein oder mehrere Nukleotide ersetzt wird.

Unter Identität zwischen zwei Nukleinsäuren wird die Identität der Nukleotide über die jeweils gesamte Nukleinsäurelänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Vector NTI Suite 7.1 Software der Firma Informax (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

Multiple alignment parameter:  
Gap opening penalty 10  
Gap extension penalty 10  
Gap separation penalty range 8  
Gap separation penalty off  
% identity for alignment delay 40

Residue specific gaps off  
Hydrophilic residue gap off  
Transition weighing 0

- 5 Pairwise alignment parameter:  
FAST algorithm on  
K-tuplesize 1  
Gap penalty 3  
Window size 5  
10 Number of best diagonals 5

15 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 1 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 1, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 2 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 2, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

25 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 3 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 3, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

30 Besonders bevorzugte EPSPS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

35 Weitere natürliche Beispiele für EPSPS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3, aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

40

Ein weiterer Gegenstand der Erfindung betrifft daher EPSPS Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 Nukleotide.

Unter "hybridisieren" versteht man die Fähigkeit eines Poly- oder Oligonukleotids, unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen vorzugsweise zu 90-100% komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze.

Die Hybridisierung erfolgt erfindungsgemäß unter stringenten Bedingungen. Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben:

Unter stringenten Hybridisierungs-Bedingungen werden insbesondere verstanden: Die über Nacht Inkubation bei 42°C in einer Lösung bestehend aus 50 % Formamid, 5 x SSC (750 mM NaCl, 75 mM Tri-Natrium Citrat), 50 mM Natrium Phosphat (pH 7,6), 5x Denhardt Lösung, 10% Dextransulfat und 20 g/ml denaturierte, gescheerte Lachsspermien-DNA, gefolgt von einem Waschen der Filter mit 0,1x SSC bei 65°C.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform A1), A2) oder A3) beschriebenen EPSPS Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ.

ID. NO. 1, 2 oder 3 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 als EPSPS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

Alle vorstehend erwähnten EPSPS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem „B-Gene Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Lycopin- $\beta$ -Cyclase, insbesondere eine chromoplastenspezifische Lycopin- $\beta$ -Cyclase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

Diese B-Gene Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen B-Gene Promotorsequenzen oder den Nukleinsäuren, kodierend eine Lycopin- $\beta$ -Cyclase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Lycopin- $\beta$ -Cyclase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Unter einer Lycopin- $\beta$ -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Lycopin in  $\gamma$ -Carotin und/oder  $\beta$ -Carotin umzuwandeln.

## Bevorzugte B-Gene Promotoren enthalten

- 5 B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder  
B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von  
Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 %  
auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder  
6 aufweist oder  
B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.  
4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder  
10 B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3)

15 Die Nukleinsäuresequenz SEQ. ID. NO. 4 stellt eine Promotorsequenz der chromoplastenspezifischen Lycopin- $\beta$ -Cyclase (B-Gene) aus *Lycopersicon esculentum* (AAZ51517) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 5 stellt eine Promotorsequenz der chromoplastenspezifischen Lycopin- $\beta$ -Cyclase (B-Gene) aus *Lycopersicon esculentum* (AAZ51521) dar.

20 Die Nukleinsäuresequenz SEQ. ID. NO. 6 stellt eine weitere Promotorsequenz der chromoplastenspezifischen Lycopin- $\beta$ -Cyclase (B-Gene) aus *Lycopersicon esculentum* dar.

25 Die Erfindung betrifft weiterhin B-Gene Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 4, 5 oder 6) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist.

30 Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße B-Gene Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 4, 5 oder 6 leicht auffinden.

35 Künstliche erfindungsgemäße B-Gene Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 4, 5 oder 6 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 4 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 4, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 5 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 5, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 6 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 6, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte B-Gene Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für B-Gene Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 4, 5 oder 6 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Ein weiterer Gegenstand der Erfindung betrifft daher B-Gene Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 4, 5 oder 6 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 Nukleotide.

Die Hybridisierungsbedingungen sind vorstehend beschrieben.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Aus-



gangssequenz.

Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform B1), B2) oder B3) beschriebenen B-Gene Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 als B-Gene Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung *Tagetes*.

Alle vorstehend erwähnten B-Gene Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), *Molecular cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem „PDS Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Phytoendesaturase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

Diese PDS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen PDS Promotorsequenzen oder den Nukleinsäuren, kodierend eine Phytoende-

saturase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Phytoendesaturase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in  
5 der Promotorsequenz.

Unter einer Phytoendesaturase wird vorzugsweise ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluen umzuwandeln.

10 Bevorzugte PDS Promotoren enthalten

C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder

C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 %  
15 auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder

C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder

C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3)  
20

Die Nukleinsäuresequenz SEQ. ID. NO. 7 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (U46919) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 8 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (X78271) dar.  
25

Die Nukleinsäuresequenz SEQ. ID. NO. 9 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* (X171023) dar.

30 Die Nukleinsäuresequenz SEQ. ID. NO. 10 stellt eine weitere Promotorsequenz der Phytoendesaturase (PDS) aus *Lycopersicon esculentum* dar.

Die Erfindung betrifft weiterhin PDS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 7, 8, 9 oder 10) durch Substitution, Insertion oder Deletion von  
35 Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße PDS Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische  
40 Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Da-

tenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 7, 8, 9 oder 10 leicht auffinden.

5 Künstliche erfindungsgemäße PDS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

10 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 7 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 7, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

15 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 8 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 8, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 9 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 9, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

25 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 10 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 10, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

30 Besonders bevorzugte PDS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

40 Weitere natürliche Beispiele für PDS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 aus verschiedenen Organismen, deren

genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

5 Ein weiterer Gegenstand der Erfindung betrifft daher PDS Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 Nukleotide.

10 Die Hybridisierungsbedingungen sind vorstehend beschrieben.

Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

15 Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

20 Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform C1), C2) oder C3) beschriebenen PDS Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12, 15, 30, 50 oder besonders bevorzugts mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 auf.

25 Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 als PDS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

30 Alle vorstehend erwähnten PDS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann  
35 beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in  
40 Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem „CHRC Promotor“ werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend ein Chromoplasten-assoziiertes Protein C regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch  
5 Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

10 Diese CHRC Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen CHRC Promotorsequenzen oder den Nukleinsäuren kodierend ein Chromoplasten-assoziiertes Protein C auffinden.  
15

Vorzugsweise werden dazu die Nukleinsäuren kodierend ein Chromoplasten-assoziiertes Protein C verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

20 Bevorzugte CHRC Promotoren enthalten

- D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
- 25 D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
- D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
- 30 D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3)

Die Nukleinsäuresequenz SEQ. ID. NO. 11 stellt eine Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke (AAV36416) dar.

35 Die Nukleinsäuresequenz SEQ. ID. NO. 12 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 13 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

40

Die Nukleinsäuresequenz SEQ. ID. NO. 14 stellt eine weitere Promotorsequenz des chromoplasten assoziiertes Protein C (CHRC) aus Gurke dar.

5 Die Erfindung betrifft weiterhin CHRC Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 11, 12, 13, oder 14) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13, oder 14 aufweist.

10 Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße CHRC Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 11, 12, 13, oder 14 leicht auffinden.

15 Künstliche erfindungsgemäße CHRC Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 11 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 11, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

25 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 12 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

30

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 13 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 13, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

35

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 14 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 14,

40

insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

5 Besonders bevorzugte CHRC Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

10 Weitere natürliche Beispiele für CHRC Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

15 Ein weiterer Gegenstand der Erfindung betrifft daher CHCRC Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 11, 12, 13, oder 14 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 Nukleotide.

20 Die Hybridisierungsbedingungen sind vorstehend beschrieben.

25 Unter einem „funktionell äquivalenten Fragment“ werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

30 Unter „im wesentlichen gleich“ wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

35 Unter „Fragmente“ werden Teilsequenzen der durch Ausführungsform D1), D2) oder D3) beschriebenen CHRC Promotoren verstanden. Vorzugsweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12, 15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 als CHRC Promotor, d.h. zur Expression von Genen in Pflanzen der

Gattung *Tagetes*.

Alle vorstehend erwähnten CHRC Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch  
5 Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-  
10 Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Mit den erfindungsgemäßen Promotoren lässt sich prinzipiell jedes Gen, also jede Nukleinsäure, kodierend ein Protein, in Pflanzen der Gattung *Tagetes* exprimieren, insbesondere blütenspezifisch exprimieren, besonders bevorzugt petalenspezifisch exprimieren.  
... mieren.

Diese in Pflanzen der Gattung *Tagetes* zu exprimierenden Gene werden im folgenden  
20 auch „Effektgene“ genannt.

Bevorzugte Effektgene sind beispielsweise Gene aus dem Biosynthesweg von Geruchsstoffen und Blütenfarben, deren Expression oder erhöhte Expression in Pflanzen der Gattung *Tagetes* zu einer Veränderung der Geruchs und/oder der Blütenfarbe von  
25 Blüten der Pflanzen der Gattung *Tagetes* führt.

Die Biosynthese von flüchtigen Geruchskomponenten, speziell in Blüten, wurde in den letzten Jahren an verschiedenen Modellorganismen wie *Clarkia breweri* und *Antirrhinum majus* L. studiert, Flüchtige Geruchskomponenten werden beispielsweise innerhalb des  
30 Monoterpen- und Phenylpropan-Stoffwechsels gebildet werden. Im ersten Fall handelt es sich um Linalool; von den Phenylpropanen sind Methyleneugenol, Benzylacetat, Methylbenzoat und Methylsalicat abgeleitet.

Für die Biosynthese von Linalool, (ISo)Methyleugenol, Benzylacetat und Methylsalicinat  
35 sind bevorzugte Gene ausgewählt aus der Gruppe Nukleinsäuren kodierend eine Linalool-Synthase (LIS), Nukleinsäuren kodierend eine S-Adenosyl-L-Met:(iso)-Eugenol-O-Methyltransferase (IEMT), Nukleinsäuren kodierend eine Acetyl-CoA-Benzylalkohol-Acetyltransferase und Nukleinsäuren kodierend eine S-Adenosyl-L-Met:Salicylsäure-Methyltransferase (SAMT). Nukleinsäuresequenzen und Proteinsequenzen zu den  
40 genannten enzymatischen Aktivitäten sind in Dudareva et al. Plant Cell 8 (1996), 1137-



1148; Wang et al. Plant Physiol. 114 (1997), 213-221 und Dudareva et al. Plant J. 14 (1998) 297-304) beschrieben.

5 Besonders bevorzugte Effektgene sind Gene aus Biosyntheswegen von biosynthetischen Produkten die in Pflanzen der Gattung Tagetes natürlicherweise, d.h. im Wildtyp oder durch genetische Veränderung des Wildtyps hergestellt werden können, insbesondere in Blüten hergestellt werden können, besonders bevorzugt in Petalen hergestellt werden können.

10 Bevorzugte biosynthetische Produkte sind Feinchemikalien.

Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und beinhaltet Verbindungen, die von einem Organismus produziert werden und in verschiedenen Industriezweigen Anwendungen finden, wie bspw., jedoch nicht beschränkt auf die pharmazeutische  
15 Industrie, die Landwirtschafts-, Kosmetik, Food und Feed-Industrie. Diese Verbindungen umfassen organische Säuren, wie beispielsweise Weinsäure, Itaconsäure und Diaminopimelinsäure, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Purin- und Pyrimidinbasen, Nukleoside und Nukleotide (wie bspw. beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology  
20 Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten), Lipide, gesättigte und ungesättigte Fettsäuren (bspw. Arachidonsäure), Diöle (bspw. Propan-1,2-diol und Butandiol), Kohlenhydrate (bspw. Hyaluronsäure und Trehalose), aromatische Verbindungen (bspw. aromatische Amine, Vanillin und Indigo), Vitamine, Carotinoide und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry,  
25 Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), Enzyme und  
30 sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien). Der Metabolismus und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.

35 1. Aminosäure-Metabolismus und Verwendungen

Die Aminosäuren umfassen die grundlegenden Struktureinheiten sämtlicher Proteine und sind somit für die normalen Zellfunktionen essentiell. Der Begriff "Aminosäure" ist im Fachgebiet bekannt. Die proteinogenen Aminosäuren, von denen es 20 Arten gibt,  
40 dienen als Struktureinheiten für Proteine, in denen sie über Peptidbindungen miteinander

der verknüpft sind, wohingegen die nicht-proteinogenen Aminosäuren (von denen Hunderte bekannt sind) gewöhnlich nicht in Proteinen vorkommen (siehe Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97 VCH: Weinheim (1985)). Die Aminosäuren können in der D- oder L-Konfiguration vorliegen, obwohl L-Aminosäuren gewöhnlich der einzige Typ sind, den man in natürlich vorkommenden Proteinen vorfindet. Biosynthese- und Abbauwege von jeder der 20 proteinogenen Aminosäuren sind sowohl bei prokaryotischen als auch eukaryotischen Zellen gut charakterisiert (siehe bspw. Stryer, L. Biochemistry, 3. Auflage, S. 578-590 (1988)). Die "essentiellen" Aminosäuren (Histidin, Isoleucin, Leucin, Lysin, Methionin, Phenylalanin, Threonin, Tryptophan und Valin), so bezeichnet, da sie aufgrund der Komplexität ihrer Biosynthese mit der Ernährung aufgenommen werden müssen, werden durch einfache Biosynthesewege in die übrigen 11 "nichtessentiellen" Aminosäuren (Alanin, Arginin, Asparagin, Aspartat, Cystein, Glutamat, Glutamin, Glycin, Prolin, Serin und Tyrosin) umgewandelt. Höhere Tiere besitzen die Fähigkeit, einige dieser Aminosäuren zu synthetisieren, jedoch müssen die essentiellen Aminosäuren mit der Nahrung aufgenommen werden, damit eine normale Proteinsynthese stattfindet.

Abgesehen von ihrer Funktion bei der Proteinbiosynthese sind diese Aminosäuren interessante Chemikalien an sich, und man hat entdeckt, daß viele bei verschiedenen Anwendungen in der Nahrungsmittel-, Futter-, Chemie-, Kosmetik-, Landwirtschafts- und pharmazeutischen Industrie zum Einsatz kommen. Lysin ist nicht nur für die Ernährung des Menschen eine wichtige Aminosäure, sondern auch für monogastrische Tiere, wie Geflügel und Schweine. Glutamat wird am häufigsten als Geschmacksadditiv (Mononatriumglutamat, MSG) sowie weithin in der Nahrungsmittelindustrie verwendet, wie auch Aspartat, Phenylalanin, Glycin und Cystein. Glycin, L-Methionin und Tryptophan werden sämtlich in der pharmazeutischen Industrie verwendet. Glutamin, Valin, Leucin, Isoleucin, Histidin, Arginin, Prolin, Serin und Alanin werden in der pharmazeutischen Industrie und der Kosmetikindustrie verwendet. Threonin, Tryptophan und D-/L-Methionin sind weitverbreitete Futtermittelzusätze (Leuchtenberger, W. (1996) Amino acids - technical production and use, S. 466-502 in Rehm et al., (Hrsg.) Biotechnology Bd. 6, Kapitel 14a, VCH: Weinheim). Man hat entdeckt, daß sich diese Aminosäuren außerdem als Vorstufen für die Synthese von synthetischen Aminosäuren und Proteinen, wie N-Acetylcystein, S-Carboxymethyl-L-cystein, (S)-5-Hydroxytryptophan und anderen, in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97, VCH, Weinheim, 1985 beschriebenen Substanzen eignen.

Die Biosynthese dieser natürlichen Aminosäuren in Organismen, die sie produzieren können, bspw. Bakterien, ist gut charakterisiert worden (für einen Überblick der bakteriellen Aminosäure-Biosynthese und ihrer Regulation, s. Umbarger, H.E. (1978) Ann. Rev. Biochem. 47: 533 - 606). Glutamat wird durch reduktive Aminierung von  $\alpha$ -

Ketoglutarat, einem Zwischenprodukt im Citronensäure-Zyklus, synthetisiert. Glutamin, Prolin und Arginin werden jeweils nacheinander aus Glutamat erzeugt. Die Biosynthese von Serin erfolgt in einem Dreischritt-Verfahren und beginnt mit 3-Phosphoglycerat (einem Zwischenprodukt bei der Glykolyse), und ergibt nach Oxidations-, Transaminierungs- und Hydrolyseschritten diese Aminosäure. Cystein und Glycin werden jeweils aus Serin produziert, und zwar die erstere durch Kondensation von Homocystein mit Serin, und die letztere durch Übertragung des Seitenketten- $\beta$ -Kohlenstoffatoms auf Tetrahydrofolat, in einer durch Serintranshydroxymethylase katalysierten Reaktion. Phenylalanin und Tyrosin werden aus den Vorstufen des Glycolyse- und Pento-

5 sephosphatweges, Erythrose-4-phosphat und Phosphoenolpyruvat in einem 9-Schritt-Biosyntheseweg synthetisiert, der sich nur in den letzten beiden Schritten nach der Synthese von Prephenat unterscheidet. Tryptophan wird ebenfalls aus diesen beiden Ausgangsmolekülen produziert, jedoch erfolgt dessen Synthese in einem 11-Schritt-

10 Weg. Tyrosin läßt sich in einer durch Phenylalaninhydroxylase katalysierten Reaktion auch aus Phenylalanin herstellen. Alanin, Valin und Leucin sind jeweils Biosyntheseprodukte aus Pyruvat, dem Endprodukt der Glykolyse. Aspartat wird aus Oxalacetat, einem Zwischenprodukt des Citratzyklus, gebildet. Asparagin, Methionin, Threonin und Lysin werden jeweils durch Umwandlung von Aspartat produziert. Isoleucin wird aus Threonin gebildet. In einem komplexen 9-Schritt-Weg erfolgt die Bildung von Histidin

15 aus 5-Phosphoribosyl-1-pyrophosphat, einem aktivierten Zucker.

20

Aminosäuren, deren Menge den Proteinbiosynthesebedarf der Zelle übersteigt, können nicht gespeichert werden, und werden stattdessen abgebaut, so daß Zwischenprodukte für die Haupt-Stoffwechselwege der Zelle bereitgestellt werden (für einen Überblick

25 siehe Stryer, L., Biochemistry, 3. Aufl. Kap. 21 "Amino Acid Degradation and the Urea Cycle"; S 495-516 (1988)). Die Zelle ist zwar in der Lage, ungewünschte Aminosäuren in nützliche Stoffwechsel-Zwischenprodukte umzuwandeln, jedoch ist die Aminosäureproduktion hinsichtlich der Energie, der Vorstufenmoleküle und der für ihre Synthese nötigen Enzyme aufwendig. Es überrascht daher nicht, daß die Aminosäure-

30 Biosynthese durch Feedback-Hemmung reguliert wird, wobei das Vorliegen einer bestimmten Aminosäure ihre eigene Produktion verlangsamt oder ganz beendet (für einen Überblick über den Rückkopplungs-Mechanismus bei Aminosäure-Biosynthesewegen, siehe Stryer, L., Biochemistry, 3. Aufl., Kap. 24, "Biosynthesis of Amino Acids and Heme", S. 575-600 (1988)). Der Ausstoß einer bestimmten Amino-

35 säure wird daher durch die Menge dieser Aminosäure in der Zelle eingeschränkt.

## *II. Vitamine, Carotinoide, Cofaktoren und Nutrazeutika-Metabolismus sowie Verwendungen*

Vitamine, Carotinoide, Cofaktoren und Nutrazeutika umfassen eine weitere Gruppe von Molekülen. Höhere Tiere haben die Fähigkeit verloren, diese zu synthetisieren und müssen sie somit aufnehmen, obwohl sie leicht durch andere Organismen, wie Bakterien, synthetisiert werden. Diese Moleküle sind entweder biologisch aktive Moleküle an sich oder Vorstufen von biologisch aktiven Substanzen, die als Elektronenträger oder Zwischenprodukte bei einer Reihe von Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Verarbeitungs-Hilfsstoffe. (Für einen Überblick über die Struktur, Aktivität und die industriellen Anwendungen dieser Verbindungen siehe bspw. Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Der Begriff "Vitamin" ist im Fachgebiet bekannt und umfaßt Nährstoffe, die von einem Organismus für eine normale Funktion benötigt werden, jedoch nicht von diesem Organismus selbst synthetisiert werden können. Die Gruppe der Vitamine kann Cofaktoren und nutrazeutische Verbindungen umfassen. Der Begriff "Cofaktor" umfaßt nicht-proteinartige Verbindungen, die für das Auftreten einer normalen Enzymaktivität nötig sind. Diese Verbindungen können organisch oder anorganisch sein; die erfindungsgemäßen Cofaktor-Moleküle sind vorzugsweise organisch. Der Begriff "Nutrazeutikum" umfaßt Nahrungsmittelzusätze, die bei Pflanzen und Tieren, insbesondere dem Menschen, gesundheitsfördernd sind. Beispiele solcher Moleküle sind Vitamine, Antioxidantien und ebenfalls bestimmte Lipide (z.B. mehrfach ungesättigte Fettsäuren).

Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Die Biosynthese dieser Moleküle in Organismen, die zu ihrer Produktion befähigt sind, wie Bakterien, ist umfassend charakterisiert worden (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996, Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for free Radical Research - Asien, abgehalten

am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press, Champaign, IL X, 374 S).

Thiamin (Vitamin B<sub>1</sub>) wird durch chemisches Kuppeln von Pyrimidin und Thiazol-Einheiten gebildet. Riboflavin (Vitamin B<sub>2</sub>) wird aus Guanosin-5'-triphosphat (GTP) und  
5 Ribose-5'-phosphat synthetisiert. Riboflavin wiederum wird zur Synthese von Flavin-mononukleotid (FMN) und Flavinadenindinukleotid (FAD) eingesetzt. Die Familie von Verbindungen, die gemeinsam als "Vitamin B6" bezeichnet werden (bspw. Pyridoxin, Pyridoxamin, Pyridoxal-5'-phosphat und das kommerziell verwendete Pyridoxin-hydrochlorid), sind alle Derivate der gemeinsamen Struktureinheit 5-Hydroxy-6-methylpyridin. Panthothenat (Pantothersäure, R-(+)-N-(2,4-Dihydroxy-3,3-dimethyl-1-oxobutyl)- $\beta$ -alanin) kann entweder durch chemische Synthese oder durch Fermentation hergestellt werden. Die letzten Schritte bei der Panthothenat-Biosynthese bestehen aus der ATP-getriebenen Kondensation von  $\beta$ -Alanin und Pantoinsäure. Die für die Biosyntheseschritte für die Umwandlung in Pantoinsäure, in  $\beta$ -Alanin und zur Kondensation in  
15 Pantothersäure verantwortlichen Enzyme sind bekannt. Die metabolisch aktive Form von Panthothenat ist Coenzym A, dessen Biosynthese über 5 enzymatische Schritte verläuft. Pantothersäure, Pyridoxal-5'-phosphat, Cystein und ATP sind die Vorstufen von Coenzym A. Diese Enzyme katalysieren nicht nur die Bildung von Panthothenat, sondern auch die Produktion von (R)-Pantoinsäure, (R)-Pantolacton, (R)-Panthenol (Provitamin B<sub>5</sub>), Pantethin (und seinen Derivaten) und Coenzym A.  
20

Die Biosynthese von Biotin aus dem Vorstufenmolekül Pimeloyl-CoA in Mikroorganismen ist ausführlich untersucht worden, und man hat mehrere der beteiligten Gene identifiziert. Es hat sich herausgestellt, daß viele der entsprechenden Proteine an der  
25 Fe-Cluster-Synthese beteiligt sind und zu der Klasse der nifS-Proteine gehören. Die Liponsäure wird von der Octanonsäure abgeleitet und dient als Coenzym beim Energie-Metabolismus, wo sie Bestandteil des Pyruvatdehydrogenasekomplexes und des  $\alpha$ -Ketoglutaratdehydrogenasekomplexes wird. Die Folate sind eine Gruppe von Substanzen, die alle von der Folsäure abgeleitet werden, die wiederum von L-Glutaminsäure, p-Aminobenzoessäure und 6-Methylpterin hergeleitet ist. Die Biosynthese der Folsäure und ihrer Derivate, ausgehend von den metabolischen Stoffwechselzwischenprodukten Guanosin-5'-triphosphat (GTP), L-Glutaminsäure und p-Aminobenzoessäure ist in bestimmten Mikroorganismen eingehend untersucht worden.  
30

35 Corrinioide (wie die Cobalamine und insbesondere Vitamin B<sub>12</sub>) und die Porphyrine gehören zu einer Gruppe von Chemikalien, die sich durch ein Tetrapyrrol-Ringsystem auszeichnen. Die Biosynthese von Vitamin B<sub>12</sub> ist hinreichend komplex, daß sie noch nicht vollständig charakterisiert worden ist, jedoch ist inzwischen ein Großteil der beteiligten Enzyme und Substrate bekannt. Nikotinsäure (Nikotinat) und Nikotinamid sind  
40 Pyridin-Derivate, die auch als "Niacin" bezeichnet werden. Niacin ist die Vorstufe der

wichtigen Coenzyme NAD (Nikotinamidadenindinukleotid) und NADP (Nikotinamidadenindinukleotidphosphat) und ihrer reduzierten Formen.

Die Produktion dieser Verbindungen im Großmaßstab beruht größtenteils auf zellfreien chemischen Synthesen, obwohl einige dieser Chemikalien ebenfalls durch großangelegte Anzucht von Mikroorganismen produziert worden sind, wie Riboflavin, Vitamin B<sub>6</sub>, Pantothenat und Biotin. Nur Vitamin B<sub>12</sub> wird aufgrund der Komplexität seiner Synthese lediglich durch Fermentation produziert. In-vitro-Verfahren erfordern einen erheblichen Aufwand an Materialien und Zeit und häufig an hohen Kosten.

### III. Purin-, Pyrimidin-, Nukleosid- und Nukleotid-Metabolismus und Verwendungen

Gene für den Purin- und Pyrimidin-Stoffwechsel und ihre entsprechenden Proteine sind wichtige Ziele für die Therapie von Tumorerkrankungen und Virusinfektionen. Der Begriff "Purin" oder "Pyrimidin" umfaßt stickstoffhaltige Basen, die Bestandteil der Nukleinsäuren, Coenzyme und Nukleotide sind. Der Begriff "Nukleotid" beinhaltet die grundlegenden Struktureinheiten der Nukleinsäuremoleküle, die eine stickstoffhaltige Base, einen Pentose-Zucker (bei RNA ist der Zucker Ribose, bei DNA ist der Zucker D-Desoxyribose) und Phosphorsäure umfassen. Der Begriff "Nukleosid" umfaßt Moleküle, die als Vorstufen von Nukleotiden dienen, die aber im Gegensatz zu den Nukleotiden keine Phosphorsäureeinheit aufweisen. Durch Hemmen der Biosynthese dieser Moleküle oder ihrer Mobilisation zur Bildung von Nukleinsäuremolekülen ist es möglich, die RNA- und DNA-Synthese zu hemmen; wird diese Aktivität zielgerichtet bei kanzerogenen Zellen gehemmt, läßt sich die Teilungs- und Replikations-Fähigkeit von Tumorzellen hemmen.

Es gibt zudem Nukleotide, die keine Nukleinsäuremoleküle bilden, jedoch als Energiespeicher (d.h. AMP) oder als Coenzyme (d.h. FAD und NAD) dienen.

Mehrere Veröffentlichungen haben die Verwendung dieser Chemikalien für diese medizinischen Indikationen beschrieben, wobei der Purin- und/oder Pyrimidin-Metabolismus beeinflusst wird (bspw. Christopherson, R.I. und Lyons, S.D. (1990) "Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents", Med. Res. Reviews 10: 505-548). Untersuchungen an Enzymen, die am Purin- und Pyrimidin-Metabolismus beteiligt sind, haben sich auf die Entwicklung neuer Medikamente konzentriert, die bspw. als Immunsuppressionsmittel oder Antiproliferantien verwendet werden können (Smith, J.L. "Enzymes in Nucleotide Synthesis" Curr. Opin. Struct. Biol. 5 (1995) 752-757; Biochem. Soc. Transact. 23 (1995) 877-902). Die Purin- und Pyrimidinbasen, Nukleoside und Nukleotide haben jedoch auch andere Einsatzmöglichkeiten: als Zwischenprodukte bei der Biosynthese verschiedener Feinchemikalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als

- kalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als Energieträger für die Zelle (bspw. ATP oder GTP) und für Chemikalien selbst, werden gewöhnlich als Geschmacksverstärker verwendet (bspw. IMP oder GMP) oder für viele medizinische Anwendungen (siehe bspw. Kuninaka, A., (1996) "Nucleotides and Related Com-  
5 pounds in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim, S. 561-612). Enzyme, die am Purin-, Pyrimidin-, Nukleosid- oder Nukleotid-Metabolismus beteiligt sind, dienen auch immer stärker als Ziele, gegen die Chemikalien für den Pflanzenschutz, einschließlich Fungiziden, Herbiziden und Insektiziden entwickelt werden.
- 10 Der Metabolismus dieser Verbindungen in Bakterien ist charakterisiert worden (für Übersichten siehe bspw. Zalkin, H. und Dixon, J.E. (1992) "De novo purin nucleotide biosynthesis" in Progress in Nucleic Acids Research and Molecular biology, Bd. 42, Academic Press, S. 259-287; und Michal, G. (1999) "Nucleotides and Nucleosides";  
15 Kap. 8 in : Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley, New York). Der Purin-Metabolismus, das Objekt intensiver Forschung, ist für das normale Funktionieren der Zelle essentiell. Ein gestörter Purin-Metabolismus in höheren Tieren kann schwere Erkrankungen verursachen, bspw. Gicht. Die Purinnukleotide werden über eine Reihe von Schritten über die Zwischenverbindung Inosin-5'-phosphat (IMP) aus Ribose-5-phosphat synthetisiert, was zur Produktion von Guanosin-5'-  
20 monophosphat (GMP) oder Adenosin-5'-monophosphat (AMP) führt, aus denen sich die als Nukleotide verwendeten Triphosphatformen leicht herstellen lassen. Diese Verbindungen werden auch als Energiespeicher verwendet, so daß ihr Abbau Energie für viele verschiedene biochemische Prozesse in der Zelle liefert. Die Pyrimidinbiosynthese erfolgt über die Bildung von Uridin-5'-monophosphat (UMP) aus Ribose-5-phosphat.  
25 UMP wiederum wird in Cytidin-5'-triphosphat (CTP) umgewandelt. Die Desoxyformen sämtlicher Nukleotide werden in einer Einschnitt-Reduktionsreaktion aus der Diphosphat-Riboseform des Nukleotides zur Diphosphat-Desoxyriboseform des Nukleotides hergestellt. Nach der Phosphorylierung können diese Moleküle an der DNA-Synthese teilnehmen.

30

#### *IV. Trehalose-Metabolismus und Verwendungen*

- Trehalose besteht aus zwei Glucosemolekülen, die über  $\alpha, \alpha$ -1,1-Bindung miteinander verknüpft sind. Sie wird gewöhnlich in der Nahrungsmittelindustrie als Süßstoff, als  
35 Additiv für getrocknete oder gefrorene Nahrungsmittel sowie in Getränken verwendet. Sie wird jedoch auch in der pharmazeutischen Industrie, der Kosmetik- und Biotechnologie-Industrie angewendet (s. bspw. Nishimoto et al., (1998) US-Patent Nr. 5 759 610; Singer, M.A. und Lindquist, S. Trends Biotech. 16 (1998) 460-467; Paiva, C.L.A. und Panek, A.D. Biotech Ann. Rev. 2 (1996) 293-314; und Shiosaka, M. J. Japan 172  
40 (1997) 97-102). Trehalose wird durch Enzyme von vielen Mikroorganismen produziert

und auf natürliche Weise in das umgebende Medium abgegeben, aus dem sie durch im Fachgebiet bekannte Verfahren gewonnen werden kann.

5 Besonders bevorzugte biosynthetische Produkte sind ausgewählt aus der Gruppe organische Säuren, Proteine, Nukleotide und Nukleoside, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Lipide und Fettsäuren, Diole, Kohlehydrate, aromatische Verbindungen, Vitamine und Cofaktoren, Enzyme und Proteine.

10 Bevorzugte organische Säuren sind Weinsäure, Itaconsäure und Diaminopimelinsäure

Bevorzugte Nukleoside und Nukleotide sind beispielsweise beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten.

15 Bevorzugte biosynthetische Produkte sind weiterhin Lipide, gesättigte und ungesättigte Fettsäuren, wie beispielsweise Arachidonsäure, Diole wie beispielsweise Propandiol und Butandiol, Kohlenhydrate, wie beispielsweise Hyaluronsäure und Trehalose, aromatische Verbindungen, wie beispielsweise aromatische Amine, Vanillin und Indigo, Vitamine und Cofaktoren, wie beispielsweise beschrieben in Ullmann's Encyclopedia of  
20 Industrial Chemistry, Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research -  
25 Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), Enzyme Polyketide (Cane *et al.* (1998) *Science* 282: 63-68), und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien.

30 Besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung *Tagetes* exprimiert werden sind demnach Gene, ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend  
35 ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Kohlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein  
40 Protein aus dem Biosyntheseweg von Carotinoiden, insbesondere Ketocarotinoiden,



Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Cofaktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen.

- 5 Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.
- 10 Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.
- 15 Ganz besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung Tagetes exprimiert werden sind demnach Gene die Proteine aus dem Biosyntheseweg von Carotinoiden kodieren.

- Insbesondere bevorzugt sind Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren
- 20 kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren
- 25 kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren
- 30 kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

- Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Keto-
- 35 Gruppe einzuführen.

- Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Canthaxanthin umzuwandeln.
- 40

Beispiele für Nukleinsäuren, kodierend eine Ketolase, und die entsprechenden Ketolasen, sind beispielsweise Sequenzen aus

- 5 Haematococcus pluvialis, insbesondere aus Haematococcus pluvialis Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),
- Haematococcus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),
- 10 Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20),
- Aliccaligenes spec. (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22),
- 15 Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).
- Synechocystis sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26).
- 20 Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 27, Protein SEQ ID NO: 28).
- 25 Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 29, Protein SEQ ID NO: 30).
- Haematococcus pluvialis  
(Accession NO: AF534876, AAN03484; Nukleinsäure: SEQ ID NO: 31, Protein : SEQ ID NO: 32)
- 30 Paracoccus sp. MBIC1143  
(Accession NO: D58420, P54972; Nukleinsäure: SEQ ID NO: 33, Protein : SEQ ID NO: 34)
- 35 Brevundimonas aurantiaca  
(Accession NO: AY166610, AAN86030; Nukleinsäure: SEQ ID NO: 35, Protein : SEQ ID NO: 36)
- 40 Nodularia spumigena NSOR10

(Accession NO: AY210783, AAO64399; Nukleinsäure: SEQ ID NO: 37, Protein : SEQ ID NO: 38)

Nostoc punctiforme ATCC 29133

- 5 (Accession NO: NZ\_AABC01000195, ZP\_00111258; Nukleinsäure: SEQ ID NO: 39, Protein : SEQ ID NO: 40)

Nostoc punctiforme ATCC 29133

- 10 (Accession NO: NZ\_AABC01000196; Nukleinsäure: SEQ ID NO: 41, Protein : SEQ ID NO: 42)

Deinococcus radiodurans R1

- 15 (Accession NO: E75561, AE001872; Nukleinsäure: SEQ ID NO: 43, Protein : SEQ ID NO: 44),

*Synechococcus* sp. WH 8102,

- 20 Nukleinsäure: Acc.-No. NZ\_AABD01000001, Basenpaar 1,354,725-1,355,528 (SEQ ID NO: 75), Protein: Acc.-No. ZP\_00115639 (SEQ ID NO: 76) (als putatives Protein annotiert),

Unter einer  $\beta$ -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, einen endständigen, linearen Rest von Lycopin in einen  $\beta$ -Ionon-Ring zu überführen.

- 25 Insbesondere wird unter einer  $\beta$ -Cyclase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\gamma$ -Carotin in  $\beta$ -Carotin umzuwandeln.

Beispiele für  $\beta$ -Cyclase-Gene sind Nukleinsäuren, kodierend eine  $\beta$ -Cyclase aus Tomate (Accession X86452). (Nukleinsäure: SEQ ID NO: 45, Protein: SEQ ID NO: 46),  
30 sowie  $\beta$ -Cyclasen der folgenden Accession Nummern:

S66350 lycopene beta-cyclase (EC 5.5.1.-) - tomato

CAA60119 lycopene synthase [*Capsicum annuum*]

S66349 lycopene beta-cyclase (EC 5.5.1.-) - common tobacco

- 35 CAA57386 lycopene cyclase [*Nicotiana tabacum*]

AAM21152 lycopene beta-cyclase [*Citrus sinensis*]

AAD38049 lycopene cyclase [*Citrus x paradisi*]

AAN86060 lycopene cyclase [*Citrus unshiu*]

AAF44700 lycopene beta-cyclase [*Citrus sinensis*]

- 40 AAK07430 lycopene beta-cyclase [*Adonis palaestina*]

	AAG10429	beta cyclase [Tagetes erecta]
	AAA81880	lycopene cyclase
	AAB53337	Lycopene beta cyclase
	AAL92175	beta-lycopene cyclase [Sandersonia aurantiaca]
5	CAA67331	lycopene cyclase [Narcissus pseudonarcissus]
	AAM45381	beta cyclase [Tagetes erecta]
	AAQ18661	lycopene beta-cyclase [Zea mays]
	AAG21133	chromoplast-specific lycopene beta-cyclase [Lycopersicon.esculentum]
	AAF18989	lycopene beta-cyclase [Daucus carota]
10	ZP_001140	hypothetical protein [Prochlorococcus marinus str. MIT9313]
	ZP_001050	hypothetical protein [Prochlorococcus marinus subsp. pastoris str. CCMP1378]
	ZP_001046	hypothetical protein [Prochlorococcus marinus subsp. pastoris str. CCMP1378]
15	ZP_001134	hypothetical protein [Prochlorococcus marinus str. MIT9313]
	ZP_001150	hypothetical protein [Synechococcus sp. WH 8102]
	AAF10377	lycopene cyclase [Deinococcus radiodurans]
	BAA29250	393aa long hypothetical protein [Pyrococcus horikoshii]
	BAC77673	lycopene beta-monocyclase [marine bacterium P99-3]
20	AAL01999	lycopene cyclase [Xanthobacter sp. Py2]
	ZP_000190	hypothetical protein [Chloroflexus aurantiacus]
	ZP_000941	hypothetical protein [Novosphingobium aromaticivorans]
	AAF78200	lycopene cyclase [Bradyrhizobium sp. ORS278]
	BAB79602	crtY [Pantoea agglomerans pv. milletiae]
25	CAA64855	lycopene cyclase [Streptomyces griseus]
	AAA21262	dycopene cyclase [Pantoea agglomerans]
	C37802	crtY protein - Erwinia uredovora
	BAB79602	crtY [Pantoea agglomerans pv. milletiae]
	AAA64980	lycopene cyclase [Pantoea agglomerans]
30	AAC44851	lycopene cyclase
	BAA09593	Lycopene cyclase [Paracoccus sp. MBIC1143]
	ZP_000941	hypothetical protein [Novosphingobium aromaticivorans]
	CAB56061	lycopene beta-cyclase [Paracoccus marcusii]
	BAA20275	lycopene cyclase [Erythrobacter longus]
35	ZP_000570	hypothetical protein [Thermobifida fusca]
	ZP_000190	hypothetical protein [Chloroflexus aurantiacus]
	AAK07430	lycopene beta-cyclase [Adonis palaestina]
	CAA67331	lycopene cyclase [Narcissus pseudonarcissus]
	AAB53337	Lycopene beta cyclase
40	BAC77673	lycopene beta-monocyclase [marine bacterium P99-3]

Eine besonders bevorzugte  $\beta$ -Cyclase ist weiterhin die chromoplastenspezifische  $\beta$ -Cyclase aus Tomate (AAG21133) (Nukleinsäure: SEQ. ID. No. 49; Protein: SEQ. ID. No. 50)

5

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

- 10 Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Zeaxanthin oder Canthaxanthin in Astaxanthin umzuwandeln.

Beispiele für ein Hydroxylase-Gene sind:

15

eine Nukleinsäure, kodierend eine Hydroxylase aus *Haematococcus pluvialis*, Accession AX038729, WO 0061764; (Nukleinsäure: SEQ ID NO: 49, Protein: SEQ ID NO: 50),

- 20 sowie Hydroxylasen der folgenden Accession Nummern:

[emb]CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108\_1, AF315289\_1, AF296158\_1, AAC49443.1, NP\_194300.1, NP\_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, AF162276\_1, AAO53295.1, AAN85601.1, CRTZ\_ERWHE, CRTZ\_PANAN, BAB79605.1, CRTZ\_ALCSP, CRTZ\_AGRAU, CAB56060.1, ZP\_00094836.1, AAC44852.1, BAC77670.1, NP\_745389.1, NP\_344225.1, NP\_849490.1, ZP\_00087019.1, NP\_503072.1, NP\_852012.1, NP\_115929.1, ZP\_00013255.1

- 30 Eine besonders bevorzugte Hydroxylase ist weiterhin die Hydroxylase aus Tomate (Accession Y14810, CrtR-b2) (Nukleinsäure: SEQ ID NO: 51; Protein: SEQ ID NO. 52).

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

35

Unter einer (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat in Isopentenylidiphosphat und Dimethylallyldiphosphate umzuwandeln.

40

Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Hydroxyethyl-ThPP und Glycerinaldehyd-3-Phosphat in 1-Deoxy-D-Xylose-5-Phosphat umzuwandeln.

- 5 Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 1-Deoxy-D-Xylose-5-Phosphat in 2-C-methyl-D-erythritol 4-Phosphat umzuwandeln

- 10 Unter einer Isopentenyl-Diphosphat- $\Delta$ - Isomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat in Dimethylallylphosphat umzuwandeln.

- 15 Unter einer Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat und Dimethylallylphosphat in Geranyl-Diphosphat umzuwandeln.

- 20 Unter einer Farnesyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, sequentiell 2 Moleküle Isopentenyl-Diphosphat mit Dimethylallyl-Diphosphat und dem resultierenden Geranyl-Diphosphat in Farnesyl-Diphosphat umzuwandeln

- 25 Unter einer Geranyl-Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Farnesyl-Diphosphat und Isopentenyl-Diphosphat in Geranyl-Geranyl-Diphosphat umzuwandeln.

Unter einer Phytoen-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Geranyl-Geranyl-Diphosphat in Phytoen umzuwandeln.

- 30 Unter einer Phytoen-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluene und/oder Phytofluene in  $\zeta$ -Carotin (Zetacarotin) umzuwandeln.

- 35 Unter einer Zeta-Carotin-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\zeta$ -Carotin in Neurosporin und/oder Neurosporin in Lycopin umzuwandeln.

Unter einem crtISO-Protein wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 7,9,7',9'-tetra-cis-Lycopin in all-trans-Lycopin umzuwandeln.

Unter einem FtsZ-Protein wird ein Protein verstanden, das eine Zellteilungs und Plastidenteilungs-fördernde Wirkung hat und Homologien zu Tubulinproteinen aufweist.

- 5 Unter einem MinD -Protein wird ein Protein verstanden, das eine multifunktionale Rolle bei der Zellteilung aufweist. Es ist eine Membran-assoziierte ATPase und kann innerhalb der Zelle eine oszillierende Bewegung von Pol zu Pol zeigen:

Beispiele für HMG-CoA-Reduktase-Gene sind:

- 10 Eine Nukleinsäure, kodierend eine HMG-CoA-Reduktase aus *Arabidopsis thaliana*, Accession NM\_106299; (Nukleinsäure: SEQ ID NO: 53, Protein: SEQ ID NO: 54),
- sowie weitere HMG-CoA-Reduktase -Gene aus anderen Organismen mit den folgenden Accession Nummern:
- 15 P54961, P54870, P54868, P54869, O02734, P22791, P54873, P54871, P23228, P13704, P54872, Q01581, P17425, P54874, P54839, P14891, P34135, O64966, P29057, P48019, P48020, P12683, P43256, Q9XEL8, P34136, O64967, P29058, P48022, Q41437, P12684, Q00583, Q9XHL5, Q41438, Q9YAS4, O76819, O28538,
- 20 Q9Y7D2, P54960, O51628, P48021, Q03163, P00347, P14773, Q12577, Q59468, P04035, O24594, P09610, Q58116, O26662, Q01237, Q01559, Q12649, O74164, O59469, P51639, Q10283, O08424, P20715, P13703, P13702, Q96UG4, Q8SQZ9, O15888, Q9TUM4, P93514, Q39628, P93081, P93080, Q944T9, Q40148, Q84MM0, Q84LS3, Q9Z9N4, Q9KLM0

25

Beispiele für (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase-Gene sind:

- Eine Nukleinsäure, kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase aus *Arabidopsis thaliana* (IyB/ISPH), ACCESSION
- 30 AY168881, (Nukleinsäure: SEQ ID NO: 55, Protein: SEQ ID NO:56),

sowie weitere (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase -Gene aus anderen Organismen mit den folgenden Accession Nummern:

- 35 T04781, AF270978\_1, NP\_485028.1, NP\_442089.1, NP\_681832.1, ZP\_00110421.1, ZP\_00071594.1, ZP\_00114706.1, ISPH\_SYNY3, ZP\_00114087.1, ZP\_00104269.1, AF398145\_1, AF398146\_1, AAD55762.1, AF514843\_1, NP\_622970.1, NP\_348471.1, NP\_562001.1, NP\_223698.1, NP\_781941.1, ZP\_00080042.1, NP\_859669.1, NP\_214191.1, ZP\_00086191.1, ISPH\_VIBCH, NP\_230334.1, NP\_742768.1,
- 40 NP\_302306.1, ISPH\_MYCLE, NP\_602581.1, ZP\_00026966.1, NP\_520563.1,

- NP\_253247.1, NP\_282047.1, ZP\_00038210.1, ZP\_00064913.1, CAA61555.1, ZP\_00125365.1, ISPH\_ACICA, EAA24703.1, ZP\_00013067.1, ZP\_00029164.1, NP\_790656.1, NP\_217899.1, NP\_641592.1, NP\_636532.1, NP\_719076.1, NP\_660497.1, NP\_422155.1, NP\_715446.1, ZP\_00090692.1, NP\_759496.1,
- 5 ISPH\_BURPS, ZP\_00129657.1, NP\_215626.1, NP\_335584.1, ZP\_00135016.1, NP\_789585.1, NP\_787770.1, NP\_769647.1, ZP\_00043336.1, NP\_242248.1, ZP\_00008555.1, NP\_246603.1, ZP\_00030951.1, NP\_670994.1, NP\_404120.1, NP\_540376.1, NP\_733653.1, NP\_697503.1, NP\_840730.1, NP\_274828.1, NP\_796916.1, ZP\_00123390.1, NP\_824386.1, NP\_737689.1, ZP\_00021222.1,
- 10 NP\_757521.1, NP\_390395.1, ZP\_00133322.1, CAD76178.1, NP\_600249.1, NP\_454660.1, NP\_712601.1, NP\_385018.1, NP\_751989.1

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene sind:

- 15 Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase aus *Lycopersicon esculentum*, ACCESSION #AF143812 (Nukleinsäure: SEQ ID NO: 57 , Protein: SEQ ID NO: 58),
- sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene aus anderen Organismen mit den folgenden Accession Nummern:
- 20 AF143812\_1, DXS\_CAPAN, CAD22530.1, AF182286\_1, NP\_193291.1, T52289, AAC49368.1, AAP14353.1, D71420, DXS\_ORYSA, AF443590\_1, BAB02345.1, CAA09804.2, NP\_850620.1, CAD22155.2, AAM65798.1, NP\_566686.1, CAD22531.1, AAC33513.1, CAC08458.1, AAG10432.1, T08140, AAP14354.1, AF428463\_1,
- 25 ZP\_00010537.1, NP\_769291.1, AAK59424.1, NP\_107784.1, NP\_697464.1, NP\_540415.1, NP\_196699.1, NP\_384986.1, ZP\_00096461.1, ZP\_00013656.1, NP\_353769.1, BAA83576.1, ZP\_00005919.1, ZP\_00006273.1, NP\_420871.1, AAM48660.1, DXS\_RHOCA, ZP\_00045608.1, ZP\_00031686.1, NP\_841218.1, ZP\_00022174.1, ZP\_00086851.1, NP\_742690.1, NP\_520342.1, ZP\_00082120.1,
- 30 NP\_790545.1, ZP\_00125266.1, CAC17468.1, NP\_252733.1, ZP\_00092466.1, NP\_439591.1, NP\_414954.1, NP\_752465.1, NP\_622918.1, NP\_286162.1, NP\_836085.1, NP\_706308.1, ZP\_00081148.1, NP\_797065.1, NP\_213598.1, NP\_245469.1, ZP\_00075029.1, NP\_455016.1, NP\_230536.1, NP\_459417.1, NP\_274863.1, NP\_283402.1, NP\_759318.1, NP\_406652.1, DXS\_SYNLE,
- 35 DXS\_SYNP7, NP\_440409.1, ZP\_00067331.1, ZP\_00122853.1, NP\_717142.1, ZP\_00104889.1, NP\_243645.1, NP\_681412.1, DXS\_SYNEL, NP\_637787.1, DXS\_CHLTE, ZP\_00129863.1, NP\_661241.1, DXS\_XANCP, NP\_470738.1, NP\_484643.1, ZP\_00108360.1, NP\_833890.1, NP\_846629.1, NP\_658213.1, NP\_642879.1, ZP\_00039479.1, ZP\_00060584.1, ZP\_00041364.1, ZP\_00117779.1,



NP\_299528.1

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene sind:

- 5 Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase aus *Arabidopsis thaliana*, ACCESSION #AF148852, (Nukleinsäure: SEQ ID NO: 59, Protein: SEQ ID NO: 60),

- 10 sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

- AF148852, AY084775, AY054682, AY050802, AY045634, AY081453, AY091405, AY098952, AJ242588, AB009053, AY202991, NP\_201085.1, T52570, AF331705\_1, BAB16915.1, AF367205\_1, AF250235\_1, CAC03581.1, CAD22156.1, AF182287\_1,
- 15 DXR\_MENPI, ZP\_00071219.1, NP\_488391.1, ZP\_00111307.1, DXR\_SYNLE, AAP56260.1, NP\_681831.1, NP\_442113.1, ZP\_00115071.1, ZP\_00105106.1, ZP\_00113484.1, NP\_833540.1, NP\_657789.1, NP\_661031.1, DXR\_BACHD, NP\_833080.1, NP\_845693.1, NP\_562610.1, NP\_623020.1, NP\_810915.1, NP\_243287.1, ZP\_00118743.1, NP\_464842.1, NP\_470690.1, ZP\_00082201.1,
- 20 NP\_781898.1, ZP\_00123667.1, NP\_348420.1, NP\_604221.1, ZP\_00053349.1, ZP\_00064941.1, NP\_246927.1, NP\_389537.1, ZP\_00102576.1, NP\_519531.1, AF124757\_19, DXR\_ZYMMO, NP\_713472.1, NP\_459225.1, NP\_454827.1, ZP\_00045738.1, NP\_743754.1, DXR\_PSEPK, ZP\_00130352.1, NP\_702530.1, NP\_841744.1, NP\_438967.1, AF514841\_1, NP\_706118.1, ZP\_00125845.1,
- 25 NP\_404661.1, NP\_285867.1, NP\_240064.1, NP\_414715.1, ZP\_00094058.1, NP\_791365.1, ZP\_00012448.1, ZP\_00015132.1, ZP\_00091545.1, NP\_629822.1, NP\_771495.1, NP\_798691.1, NP\_231885.1, NP\_252340.1, ZP\_00022353.1, NP\_355549.1, NP\_420724.1, ZP\_00085169.1, EAA17616.1, NP\_273242.1, NP\_219574.1, NP\_387094.1, NP\_296721.1, ZP\_00004209.1, NP\_823739.1,
- 30 NP\_282934.1, BAA77848.1, NP\_660577.1, NP\_760741.1, NP\_641750.1, NP\_636741.1, NP\_829309.1, NP\_298338.1, NP\_444964.1, NP\_717246.1, NP\_224545.1, ZP\_00038451.1, DXR\_KITGR, NP\_778563.1.

Beispiele für Isopentenyl-Diphosphat- $\Delta$ -Isomerase-Gene sind:

- 35 Eine Nukleinsäure, kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase aus *Adonis palaestina* clone ApiPI28, (ipiAa1), ACCESSION #AF188060, veröffentlicht durch Cunningham, F.X. Jr. and Gantt, E.: Identification of multi-gene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli*, *Plant Cell Physiol.* 41 (1), 119-123 (2000) (Nukleinsäure: SEQ ID NO: 61; Protein:
- 40

SEQ ID NO: 62),

sowie weitere Isopentenyl-Diphosphat- $\Delta$ -Isomerase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

5

Q38929, O48964, Q39472, Q13907, O35586, P58044, O42641, O35760, Q10132, P15496, Q9YB30, Q8YNH4, Q42553, O27997, P50740, O51627, O48965, Q8KFR5, Q39471, Q39664, Q9RVE2, Q01335, Q9HHE4, Q9BXS1, Q9KWF6, Q9CIF5, Q88WB6, Q92BX2, Q8Y7A5, Q8TT35, Q9KK75, Q8NN99, Q8XD58, Q8FE75,

10

Q46822, Q9HP40, P72002, P26173, Q9Z5D3, Q8Z3X9, Q8ZM82, Q9X7Q6, O13504, Q9HFW8, Q8NJL9, Q9UUQ1, Q9NH02, Q9M6K9, Q9M6K5, Q9FXR6, O81691, Q9S7C4, Q8S3L8, Q9M592, Q9M6K3, Q9M6K7, Q9FV48, Q9LLB6, Q9AVJ1, Q9AVG8, Q9M6K6, Q9AVJ5, Q9M6K2, Q9AYS5, Q9M6K8, Q9AVG7, Q8S3L7, Q8W250, Q94IE1, Q9AVI8, Q9AYS6, Q9SAY0, Q9M6K4, Q8GVZ0, Q84RZ8,

15

Q8KZ12, Q8KZ66, Q8FND7, Q88QC9, Q8BFZ6, BAC26382, CAD94476.

Beispiele für Geranyl-Diphosphat-Synthase -Gene sind:

20

Eine Nukleinsäure, kodierend eine Geranyl-Diphosphat-Synthase aus *Arabidopsis thaliana*, ACCESSION #Y17376, Bouvier, F., Suire, C., d'Harlingue, A., Backhaus, R. A. and Camara, B.; Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells, Plant J. 24 (2), 241-252 (2000) (Nukleinsäure: SEQ ID NO: 63, Protein: SEQ ID NO: 64),

25

sowie weitere Geranyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

Q9FT89, Q8LKJ2, Q9FSW8, Q8LKJ3, Q9SBR3, Q9SBR4, Q9FET8, Q8LKJ1, Q84LG1, Q9JK86

30

Beispiele für Farnesyl-Diphosphat-Synthase-Gene sind:

35

Eine Nukleinsäure, kodierend eine Farnesyl-Diphosphat-Synthase aus *Arabidopsis thaliana* (FPS1), ACCESSION #U80605, veröffentlicht durch Cunillera, N., Arro, M., Delourme, D., Karst, F., Boronat, A. und Ferrer, A.: *Arabidopsis thaliana* contains two differentially expressed farnesyl-diphosphate synthase genes, J. Biol. Chem. 271 (13), 7774-7780 (1996), (Nukleinsäure: SEQ ID NO: 65, Protein: SEQ ID NO: 66),

sowie weitere Farnesyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

5 P53799, P37268, Q02769, Q09152, P49351, O24241, Q43315, P49352, O24242, P49350, P08836, P14324, P49349, P08524, O66952, Q08291, P54383, Q45220, P57537, Q8K9A0, P22939, P45204, O66126, P55539, Q9SWH9, Q9AVI7, Q9FRX2, Q9AYS7, Q94IE8, Q9FXR9, Q9ZWF6, Q9FXR8, Q9AR37, O50009, Q94IE9, Q8RVK7, Q8RVQ7, O04882, Q93RA8, Q93RB0, Q93RB4, Q93RB5, Q93RB3, Q93RB1, Q93RB2, Q920E5.

10

Beispiele für Geranyl-geranyl-Diphosphat-Synthase -Gene sind:

15 Eine Nukleinsäure, kodierend eine Geranyl-geranyl-Diphosphat-Synthase aus *Sinaps alba*, ACCESSION #X98795, veröffentlicht durch Bonk, M., Hoffmann, B., Von Lintig, J., Schledz, M., Al-Babili, S., Hobeika, E., Kleinig, H. and Beyer, P.: Chloroplast import of four carotenoid biosynthetic enzymes in vitro reveals differential fates prior to membrane binding and oligomeric assembly, *Eur. J. Biochem.* 247 (3), 942-950 (1997), (Nukleinsäure: SEQ ID NO: 67, Protein: SEQ ID NO: 68),

20 sowie weitere Geranyl-geranyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

25 P22873, P34802, P56966, P80042, Q42698, Q92236, O95749, Q9WTN0, Q50727, P24322, P39464, Q9FXR3, Q9AYN2, Q9FXR2, Q9AVG6, Q9FRW4, Q9SXZ5, Q9AVJ7, Q9AYN1, Q9AVJ4, Q9FXR7, Q8LSC5, Q9AVJ6, Q8LSC4, Q9AVJ3, Q9SSU0, Q9SXZ6, Q9SST9, Q9AVJ0, Q9AVI9, Q9FRW3, Q9FXR5, Q94IF0, Q9FRX1, Q9K567, Q93RA9, Q93QX8, CAD95619, EAA31459

Beispiele für Phytoen-Synthase-Gene sind:

30

35 Eine Nukleinsäure, kodierend eine Phytoen-Synthase aus *Erwinia uredovora*, ACCESSION # D90087; veröffentlicht durch Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., Nakamura, K. und Harashima, K.: Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*; *J. Bacteriol.* 172 (12), 6704-6712 (1990), (Nukleinsäure: SEQ ID NO: 69, Protein: SEQ ID NO: 70),

sowie weitere Phytoen-Synthase -Gene aus anderen Organismen mit den folgenden Accession Nummern:

40

CAB39693, BAC69364, AAF10440, CAA45350, BAA20384, AAM72615, BAC09112, CAA48922, P\_001091, CAB84588, AAF41518, CAA48155, AAD38051, AAF33237, AAG10427, AAA34187, BAB73532, CAC19567, AAM62787, CAA55391, AAB65697, AAM45379, CAC27383, AAA32836, AAK07735, BAA84763, P\_000205, AAB60314, 5 P\_001163, P\_000718, AAB71428, AAA34153, AAK07734, CAA42969, CAD76176, CAA68575, P\_000130, P\_001142, CAA47625, CAA85775, BAC14416, CAA79957, BAC76563, P\_000242, P\_000551, AAL02001, AAK15621, CAB94795, AAA91951, P\_000448

10 Beispiele für Phytoen-Desaturase-Gene sind:

Eine Nukleinsäure, kodierend eine Phytoen-Desaturase aus *Erwinia uredovora*, AC-  
 15 CSSION # D90087; veröffentlicht durch Misawa, N., Nakagawa, M., Kobayashi, K.,  
 Yamano, S., Izawa, Y., Nakamura, K. und Harashima, K.: Elucidation of the *Erwinia ure-*  
*dovora* carotenoid biosynthetic pathway by functional analysis of gene products ex-  
 pressed in *Escherichia coli*; J. Bacteriol. 172 (12), 6704-6712 (1990), (Nukleinsäure:  
 SEQ ID NO: 71, Protein: SEQ ID NO: 72),

20 sowie weitere Phytoen-Desaturase -Gene aus anderen Organismen mit den folgenden  
 Accession Nummern:

AAL15300, A39597, CAA42573, AAK51545, BAB08179, CAA48195, BAB82461,  
 AAK92625, CAA55392, AAG10426, AAD02489, AAO24235, AAC12846, AAA99519,  
 AAL38046, CAA60479, CAA75094, ZP\_001041, ZP\_001163, CAA39004, CAA44452,  
 25 ZP\_001142, ZP\_000718, BAB82462, AAM45380, CAB56040, ZP\_001091, BAC09113,  
 AAP79175, AAL80005, AAM72642, AAM72043, ZP\_000745, ZP\_001141, BAC07889,  
 CAD55814, ZP\_001041, CAD27442, CAE00192, ZP\_001163, ZP\_000197, BAA18400,  
 AAG10425, ZP\_001119, AAF13698, 2121278A, AAB35386, AAD02462, BAB68552,  
 CAC85667, AAK51557, CAA12062, AAG51402, AAM63349, AAF85796, BAB74081,  
 30 AAA91161, CAB56041, AAC48983, AAG14399, CAB65434, BAB73487, ZP\_001117,  
 ZP\_000448, CAB39695, CAD76175, BAC69363, BAA17934, ZP\_000171, AAF65586,  
 ZP\_000748, BAC07074, ZP\_001133, CAA64853, BAB74484, ZP\_001156, AAF23289,  
 AAG28703, AAP09348, AAM71569, BAB69140, ZP\_000130, AAF41516, AAG18866,  
 CAD95940, NP\_656310, AAG10645, ZP\_000276, ZP\_000192, ZP\_000186,  
 35 AAM94364, EAA31371, ZP\_000612, BAC75676, AAF65582

Beispiele für Zeta-Carotin-Desaturase-Gene sind:

40 Eine Nukleinsäure, kodierend eine Zeta-Carotin-Desaturase aus *Narcissus pseudonar-*  
*cissus*, ACCESSION #AJ224683, veröffentlicht durch Al-Babili, S., Oelschlegel, J. and

Beyer, P.: A cDNA encoding for beta carotene desaturase (Accession No. AJ224683) from *Narcissus pseudonarcissus* L.. (PGR98-103), *Plant Physiol.* 117, 719-719 (1998), (Nukleinsäure: SEQ ID NO: 73, Protein: SEQ ID NO: 74),

- 5 sowie weitere Zeta-Carotin-Desaturase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

Q9R6X4, Q38893, Q9SMJ3, Q9SE20, Q9ZTP4, O49901, P74306, Q9FV46, Q9RCT2, ZDS\_NARPS, BAB68552.1, CAC85667.1, AF372617\_1, ZDS\_TARER, CAD55814.1, 10 CAD27442.1, 2121278A, ZDS\_CAPAN, ZDS\_LYCES, NP\_187138.1, AAM63349.1, ZDS\_ARATH, AAA91161.1, ZDS\_MAIZE, AAG14399.1, NP\_441720.1, NP\_486422.1, ZP\_00111920.1, CAB56041.1, ZP\_00074512.1, ZP\_00116357.1, NP\_681127.1, ZP\_00114185.1, ZP\_00104126.1, CAB65434.1, NP\_662300.1

- 15 Beispiele für crtISO-Gene sind:

Eine Nukleinsäure, kodierend eine crtISO aus *Lycopersicon esculentum*; ACCESSION #AF416727, veröffentlicht durch Isaacson, T., Ronen, G., Zamir, D. and Hirschberg, J.: Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the pro- 20 duction of beta-carotene and xanthophylls in plants; *Plant Cell* 14 (2), 333-342 (2002), (Nukleinsäure: SEQ ID NO: 75, Protein: SEQ ID NO: 76),

sowie weitere crtISO -Gene aus anderen Organismen mit den folgenden Accession Nummern:

25

AAM53952

Beispiele für FtsZ-Gene sind:

30 Eine Nukleinsäure, kodierend eine FtsZ aus *Tagetes erecta*, ACCESSION #AF251346, veröffentlicht durch Moehs, C.P., Tian, L., Osteryoung, K.W. and Dellapenna, D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development *Plant Mol. Biol.* 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 77, Protein: SEQ ID NO: 78),

35

sowie weitere FtsZ -Gene aus anderen Organismen mit den folgenden Accession Nummern:

40 CAB89286.1, AF205858\_1, NP\_200339.1, CAB89287.1, CAB41987.1, AAA82068.1, T06774, AF383876\_1, BAC57986.1, CAD22047.1, BAB91150.1, ZP\_00072546.1,

- NP\_440816.1, T51092, NP\_683172.1, BAA85116.1, NP\_487898.1, JC4289, BAA82871.1, NP\_781763.1, BAC57987.1, ZP\_00111461.1, T51088, NP\_190843.1, ZP\_00060035.1, NP\_846285.1, AAL07180.1, NP\_243424.1, NP\_833626.1, AAN04561.1, AAN04557.1, CAD22048.1, T51089, NP\_692394.1, NP\_623237.1,
- 5 NP\_565839.1, T51090, CAA07676.1, NP\_113397.1, T51087, CAC44257.1, E84778, ZP\_00105267.1, BAA82091.1, ZP\_00112790.1, BAA96782.1, NP\_348319.1, NP\_471472.1, ZP\_00115870.1, NP\_465556.1, NP\_389412.1, BAA82090.1, NP\_562681.1, AAM22891.1, NP\_371710.1, NP\_764416.1, CAB95028.1, FTSZ\_STRGR, AF120117\_1, NP\_827300.1, JE0282, NP\_626341.1, AAC45639.1,
- 10 NP\_785689.1, NP\_336679.1, NP\_738660.1, ZP\_00057764.1, AAC32265.1, NP\_814733.1, FTSZ\_MYCKA, NP\_216666.1, CAA75616.1, NP\_301700.1, NP\_601357.1, ZP\_00046269.1, CAA70158.1, ZP\_00037834.1, NP\_268026.1, FTSZ\_ENTHR, NP\_787643.1, NP\_346105.1, AAC32264.1, JC5548, AAC95440.1, NP\_710793.1, NP\_687509.1, NP\_269594.1, AAC32266.1, NP\_720988.1,
- 15 NP\_657875.1, ZP\_00094865.1, ZP\_00080499.1, ZP\_00043589.1, JC7087, NP\_660559.1, AAC46069.1, AF179611\_14, AAC44223.1, NP\_404201.1.

Beispiele für MinD -Gene sind:

- 20 Eine Nukleinsäure, kodierend eine MinD aus *Tagetes erecta*, ACCESSION #AF251019, veröffentlicht durch Moehs, C.P., Tian, L., Osteryoung, K.W. und Dellapenna, D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development; Plant Mol. Biol. 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 79, Protein: SEQ ID NO: 80),

25

sowie weitere MinD -Gene mit den folgenden Accession Nummern:

- NP\_197790.1, BAA90628.1, NP\_038435.1, NP\_045875.1, AAN33031.1, NP\_050910.1, CAB53105.1, NP\_050687.1, NP\_682807.1, NP\_487496.1,
- 30 ZP\_00111708.1, ZP\_00071109.1, NP\_442592.1, NP\_603083.1, NP\_782631.1, ZP\_00097367.1, ZP\_00104319.1, NP\_294476.1, NP\_622555.1, NP\_563054.1, NP\_347881.1, ZP\_00113908.1, NP\_834154.1, NP\_658480.1, ZP\_00059858.1, NP\_470915.1, NP\_243893.1, NP\_465069.1, ZP\_00116155.1, NP\_390677.1, NP\_692970.1, NP\_298610.1, NP\_207129.1, ZP\_00038874.1, NP\_778791.1,
- 35 NP\_223033.1, NP\_641561.1, NP\_636499.1, ZP\_00088714.1, NP\_213595.1, NP\_743889.1, NP\_231594.1, ZP\_00085067.1, NP\_797252.1, ZP\_00136593.1, NP\_251934.1, NP\_405629.1, NP\_759144.1, ZP\_00102939.1, NP\_793645.1, NP\_699517.1, NP\_460771.1, NP\_860754.1, NP\_456322.1, NP\_718163.1, NP\_229666.1, NP\_357356.1, NP\_541904.1, NP\_287414.1, NP\_660660.1,
- 40 ZP\_00128273.1, NP\_103411.1, NP\_785789.1, NP\_715361.1, AF149810\_1,

NP\_841854.1, NP\_437893.1, ZP\_00022726.1, EAA24844.1, ZP\_00029547.1, NP\_521484.1, NP\_240148.1, NP\_770852.1, AF345908\_2, NP\_777923.1, ZP\_00048879.1, NP\_579340.1, NP\_143455.1, NP\_126254.1, NP\_142573.1, NP\_613505.1, NP\_127112.1, NP\_712786.1, NP\_578214.1, NP\_069530.1, 5 NP\_247526.1, AAA85593.1, NP\_212403.1, NP\_782258.1, ZP\_00058694.1, NP\_247137.1, NP\_219149.1, NP\_276946.1, NP\_614522.1, ZP\_00019288.1, CAD78330.1

10 Die Erfindung betrifft ferner eine genetisch veränderte Pflanze der Gattung Tagetes, wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch die erfindungsgemäßen Promotoren.

15 Wie vorstehend erwähnt, wird unter „Expressionsaktivität“ erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate verstanden.

20 Unter „spezifischer Expressionsaktivität“ wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

25 Bei einer „verursachten Expressionsaktivität“ oder „verursachten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp der Pflanze der Gattung Tagetes nicht vorhanden war.

30 Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes kein Ketolase-Gen auf. Die Regulation der Expression des Ketolase-Gens in der Pflanze durch die erfindungsgemäßen Promotoren führt somit zu einer Verursachung der Expressionsrate.

Bei einer „erhöhten Expressionsaktivität“ oder „erhöhten Expressionsrate“ im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp der Pflanze der Gattung Tagetes in einer bestimmten Zeit die gebildete Menge des Proteins erhöht.

35 Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes ein Hydroxylase-Gen auf. Die Regulation der Expression des Hydroxylase-Gens in der Pflanze durch die erfindungsgemäßen Promotoren führt somit zu einer Erhöhung der Expressionsrate.

40 In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes wird die Regulation der Expression von Genen in der

Pflanze durch die erfindungsgemäßen Promotoren dadurch erreicht, dass man

- 5 a) einen oder mehrere erfindungsgemäße Promotoren in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer endogenen Gene unter der Kontrolle der eingebrachten erfindungsgemäßen Promotoren erfolgt oder
- 10 b) ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen, erfindungsgemäßen Promotoren erfolgt oder
- 15 c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungsgemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Gene in die Pflanze einbringt.
- 20 In einer bevorzugten Ausführungsform bringt man gemäß Merkmal c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungsgemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Gene, in die Pflanze ein. Die Integration der Nukleinsäurekonstrukte in der Pflanze der Gattung *Tagetes* kann dabei intrachromosomal oder extrachromosomal erfolgen.
- 25 Bevorzugte erfindungsgemäße Promotoren und bevorzugte zu exprimierende Gene (Effektgene) sind vorstehend beschrieben.
- Im folgenden wird exemplarisch die Herstellung der genetisch veränderten Pflanzen der Gattung *Tagetes* mit erhöhter oder verursachter Expressionsrate eines Effektgens beschrieben.
- Die Transformation kann bei den Kombinationen von genetischen Veränderungen einzeln oder durch Mehrfachkonstrukte erfolgen.
- 30 Die Herstellung der transgenen Pflanzen erfolgt vorzugsweise durch Transformation der Ausgangspflanzen, mit einem Nukleinsäurekonstrukt, das mindestens einen der vorstehend beschriebenen, erfindungsgemäßen Promotoren enthält, die mit einem zu exprimierenden Effektgen und gegebenenfalls weiteren Regulationssignalen funktionell verknüpft sind.
- 35 Diese Nukleinsäurekonstrukte, in denen die erfindungsgemäßen Promotoren und Effektgene funktionell verknüpft sind, werden im folgenden auch Expressionskassetten genannt.



- Die Expressionskassetten können weitere Regulationssignale enthalten, also regulative Nukleinsäuresequenzen, welche die Expression der Effektgene in der Wirtszelle steuern. Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der kodierenden Sequenz, mindestens einen erfindungsgemäßen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungssignal und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwischenliegenden kodierenden Sequenz des Effektgens für mindestens eines der vorstehend beschriebenen Gene operativ verknüpft sind.
- 10 Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, das jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann.
- 15 Im folgenden werden beispielhaft die bevorzugten Nukleinsäurekonstrukte, Expressionskassetten und Vektoren für Pflanzen und Verfahren zur Herstellung von transgenen Pflanzen, sowie die transgenen Pflanzen der Gattung *Tagetes* selbst beschrieben.
- 20 Die zur operativen Verknüpfung bevorzugten, aber nicht darauf beschränkten Sequenzen, sind Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten und Translationsverstärkern wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693 -8711).
- 25 Die Herstellung einer Expressionskassette erfolgt vorzugsweise durch Fusion mindestens eines erfindungsgemäßen Promotors mit mindestens einem Gen, vorzugsweise mit einem der vorstehend beschriebenen Effektgene, und vorzugsweise einer zwischen Promotor und Nukleinsäure-Sequenz inserierten Nukleinsäure, die für ein plastidenspezifisches Transitpeptid kodiert, sowie einem Polyadenylierungssignal nach gängigen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley-Interscience (1987), beschrieben sind.
- 30
- 35

Die vorzugsweise insertierte Nukleinsäuren, kodierend ein plastidäres Transitpeptid, gewährleisten die Lokalisation in Plastiden und insbesondere in Chromoplasten.

Es können auch Expressionskassetten verwendet werden, deren Nukleinsäure-Sequenz für ein Effektgen-Produkt-Fusionsprotein kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das die Translokation des Polypeptides steuert. Bevorzugt sind für die Chromoplasten spezifische Transitpeptide, welche nach Translokation der Effektgene in die Chromoplasten vom Effektgenprodukt-Teil enzymatisch abgespalten werden.

Insbesondere bevorzugt ist das Transitpeptid, das von der plastidären *Nicotiana tabacum* Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco (rbcS) oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2 oder dessen funktionellem Äquivalent abgeleitet ist.

Besonders bevorzugt sind Nukleinsäure-Sequenzen von drei Kassetten des Plastiden-Transitpeptids der plastidären Transketolase aus Tabak in drei Leserastern als KpnI/BamHI Fragmente mit einem ATG-Codon in der NcoI Schnittstelle:

20 pTP09

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCTCACTCTCTCTCAAGCTATCCTCTCTC  
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
25 CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACTCCCGCCGCCG-  
TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGC-  
GATTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAAGTGAAGTGCCTG-  
GATCC\_BamHI

30 pTP10

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCTCACTCTCTCTCAAGCTATCCTCTCTC  
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACTCCCGCCGCCG-  
35 TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGC-  
GATTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAAGTGAAGTGCCTG-  
GATCC\_BamHI

40 pTP11

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCTCACTCTCTCTCAAGCTATCCTCTCTC  
 GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCT-  
 CACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACTCCCGCCGCCG-  
 TACTCCTTCCTCCGCCGCCGCCGCCGCGTCTGTAAGGTCACCGGC-  
 5 GATTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAAGTCTGAGACTGCGGG-  
 GATCC\_BamHI

Weitere Beispiele für ein plastidäres Transitpeptid sind das Transitpeptid der plastidären Isopentenyl-pyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* und das  
 10 Transitpeptid der kleinen Untereinheit der Ribulosebiphosphat Carboxylase (rbcS) aus Erbse (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. Nucl. Acids Res. 16: 11380).

Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich  
 15 gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen.

Bevorzugt sind, wie vorstehend beschrieben, synthetische Nukleotid-Sequenzen mit  
 20 Kodons, die von Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können aus Kodons mit der höchsten Proteinhäufigkeit bestimmt werden, die in den meisten interessanten Pflanzenspezies exprimiert werden.

Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente  
 25 manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adaptern oder Linker angesetzt werden.

Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allgemeinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger  
 30 als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zur Wirtspflanze sein. Die Expressionskassette beinhaltet vorzugsweise in der 5'-3'-Transkriptionsrichtung den Promotor, eine kodierende Nukleinsäuresequenz oder ein Nukleinsäurekonstrukt und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche

sind gegeneinander beliebig austauschbar.

Beispiele für einen Terminator sind der 35S-Terminator (Guerineau et al. (1988) Nucl Acids Res. 16: 11380), der nos Terminator (Depicker A, Stachel S; Dhaese P, 5  
Zambryski P, Goodman HM. Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1982;1(6):561-73) oder der ocs Terminator (Gielen, J, de Beuckeleer, M, Seurinck, J, Debroek, H, de Greve, H, Lemmers, M, van Montagu, M, Schell, J (1984) The complete sequence of the TL-DNA of the Agrobacterium tumefaciens plasmid pTiAch5. EMBO J. 3: 835-846).

10 Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt werden. Wo Insertionen, Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können *in vitro*-Mutagenese, "primer-repair", Restriktion 15  
oder Ligation verwendet werden.

Bei geeigneten Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "bluntends", können komplementäre Enden der Fragmente für die Li- 20  
gation zur Verfügung gestellt werden.

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus Agrobacterium tumefaciens, insbesondere des Gens 3 der T-DNA (Octopine Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente. 25

Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet.

30 Dazu können an sich bekannte Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengewebe oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt werden.

Geeignete Methoden zur Transformation von Pflanzen sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone – die sogenannte "particle bombardment" Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der, vorstehend beschriebene, durch *Agrobacterium* vermittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene 35  
Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von 40

S.D. Kung und R. Wu, Academic Press (1993), 128-143 sowie in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225) beschrieben.

5 Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, *Agrobacterium tumefaciens* zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) oder besonders bevorzugt pSUN2, pSUN3, pSUN4 oder pSUN5 (WO 02/00900).

10 Mit einem Expressionsplasmid transformierte Agrobakterien können in bekannter Weise zur Transformation von Pflanzen verwendet werden, z.B. indem verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

15 Zur bevorzugten Herstellung von genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, wird die fusionierte Expressionskassette in einen Vektor, beispielsweise pBin19 oder insbesondere pSUN5 und pSUN3 kloniert, der geeignet ist, in *Agrobacterium tumefaciens* transformiert zu werden. Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen verwendet werden, indem beispielsweise  
20 weise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

Die Transformation von Pflanzen durch Agrobakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die ein oder mehrere in die Expressionskassette integrierte Gene enthalten.

30 Zur Transformation einer Wirtspflanze mit einem oder mehreren erfindungsgemäßen Effektgenen wird eine Expressionskassette als Insertion in einen rekombinanten Vektor eingebaut, dessen Vektor-DNA zusätzliche funktionelle Regulationssignale, beispielsweise Sequenzen für Replikation oder Integration enthält. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology"  
35 (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrieben.

Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Expressionskassetten in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a.  
40 pJIT117 (Guerineau et al. (1988) Nucl. Acids Res. 16 :11380), pBR332, pUC-Serien,

M13mp-Serien und pACYC184. Besonders geeignet sind binäre Vektoren, die sowohl in *E. coli* als auch in Agrobakterien replizieren können.

5 Die Erfindung betrifft daher weiterhin eine genetisch veränderte Pflanze der Gattung Tagetes, enthaltend einen erfindungsgemäßen Promotor und funktionell verknüpft ein zu exprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

10 Bevorzugte erfindungsgemäße Promotoren und bevorzugte Effektgene sind vorstehend beschrieben.

Insbesondere bevorzugt sind Effektgene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

30 Bevorzugte, genetisch veränderte Pflanzen der Gattung Tagetes sind Marigold, Tagetes erecta, Tagetes patula, Tagetes lucida, Tagetes pringlei, Tagetes palmeri, Tagetes minuta oder Tagetes campanulata.

35 Durch die erfindungsgemäßen Promotoren ist es mit Hilfe der vorstehend beschriebenen, erfindungsgemäßen Verfahren möglich, in den vorstehend beschriebenen, erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes die Stoffwechselwege zu spezifischen biosynthetischen Produkten zu regulieren.

40 Dazu werden beispielsweise Stoffwechselwege, die zu einem spezifischen biosynthetischen Produkt führen, durch Verursachung oder Erhöhung der Transkriptionsrate bzw. Expressionsrate von Genen dieses Biosyntheseweges verstärkt, indem die erhöhte

Proteinmenge zu einer erhöhten Gesamttätigkeit dieser Proteine des gewünschten Biosyntheseweges und damit durch einem verstärkten Stoffwechselfluß zu dem gewünschten biosynthetischen Produkt führt.

- 5 Je nach gewünschtem biosynthetischen Produkt muss die Transkriptionsrate bzw. Expressionsrate verschiedener Gene erhöht bzw. reduziert werden. In der Regel ist es vorteilhaft, die Transkriptionsrate bzw. Expressionsrate mehrere Gene zu verändern, d.h. die Transkriptionsrate bzw. Expressionsrate einer Kombination von Gene zu Erhöhen und/oder die Transkriptionsrate bzw. Expressionsrate einer Kombination von
- 10 Gene zu reduzieren.

In den erfindungsgemäßen genetisch veränderten Pflanzen ist mindestens eine erhöhte oder verursachte Expressionsrate eines Gens auf einen erfindungsgemäßen Promotor zurückzuführen.

- 15 Weitere, zusätzliche veränderte, d.h. zusätzlich erhöhte oder zusätzlich reduzierte Expressionsraten von weiteren Genen in genetisch veränderten Pflanzen können, müssen aber nicht auf die erfindungsgemäßen Promotoren zurück gehen.

- 20 Die Erfindung betrifft daher ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von erfindungsgemäßen, genetisch veränderten Pflanzen der Gattung Tagetes.

- Die Erfindung betrifft insbesondere ein Verfahren zur Herstellung von Carotinoiden durch Kultivierung von erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein
- 25
- 30
- 35

FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Die Carotinoide sind vorzugsweise ausgewählt aus der Gruppe Phytoen, Phytofluen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Insbesondere betrifft die Erfindung weiterhin ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Die Ketocarotinoide sind vorzugsweise ausgewählt aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Im erfindungsgemäßen Verfahren zur Herstellung von biosynthetischen Produkten, insbesondere Carotinoiden, vorzugsweise Ketocarotinoiden, wird vorzugsweise dem Kultivierungsschritt der genetisch veränderten Pflanzen ein Ernten der Pflanzen und ein Isolieren der biosynthetischen Produkte, insbesondere Carotinoide, vorzugsweise Ketocarotinoide aus den Pflanzen, vorzugsweise aus den Petalen der Pflanzen, angeschlossen.

Die genetisch veränderten Pflanzen der Gattung Tagetes werden in an sich bekannter Weise auf Nährböden gezogen und entsprechend geerntet.

Die Isolierung von Ketocarotinoiden aus den geernteten Blütenblättern erfolgt beispielsweise in an sich bekannter Weise, beispielsweise durch Trocknung und an-



5 schließender Extraktion und gegebenenfalls weiterer chemischer oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie. Die Isolierung von Ketocarotinoiden aus den Blütenblättern erfolgt beispielsweise bevorzugt durch organische Lösungsmittel wie Aceton, Hexan, Heptan, Ether oder tert.-Methylbutylether.

10 Weitere Isolierverfahren von Ketocarotinoiden, insbesondere aus Blütenblättern, sind beispielsweise in Egger und Kleinig (Phytochemistry (1967) 6, 437-440) und Egger (Phytochemistry (1965) 4, 609-618) beschrieben.

Ein besonders bevorzugtes Ketocarotinoid ist Astaxanthin.

15 Die Ketocarotinoide fallen im erfindungsgemäßen Verfahren in Blütenblättern in Form ihrer Mono- oder Diester mit Fettsäuren an. Einige nachgewiesene Fettsäuren sind z.B. Myristinsäure, Palmitinsäure, Stearinsäure, Ölsäure, Linolensäure, und Laurinsäure (Kamata und Simpson (1987) Comp. Biochem. Physiol. Vol. 86B(3), 587-591).

20 Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen oder Pflanzenteile, wie insbesondere Blütenblätter mit erhöhtem Gehalt an biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin, können auch beispielsweise direkt oder nach an sich bekannter Prozessierung als Nahrungsmittel oder Futtermittel oder als Futter- und Nahrungsergänzungsmittel verwendet werden.

25 Ferner können die genetisch veränderten Pflanzen zur Herstellung von biosynthetischen Produkt-, insbesondere Carotinoid-, insbesondere Ketocarotinoid-, insbesondere Astaxanthin-haltigen Extrakten und/oder zur Herstellung von Futter- und Nahrungsergänzungsmitteln sowie von Kosmetika und Pharmazeutika verwendet werden.

30 Die genetisch veränderten Pflanzen der Gattung Tagetes weisen im Vergleich zum Wildtyp einen erhöhten Gehalt an dem gewünschten biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin auf.

35 Unter einem erhöhten Gehalt wird in diesem Fall auch ein verursachter Gehalt an Ketocarotinoiden, bzw. Astaxanthin verstanden.

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

40

Allgemeine Experimentelle Bedingungen:  
Sequenzanalyse rekombinanter DNA

Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma Licor (Vertrieb durch MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Beispiel 1:

Amplifikation einer DNA, die die gesamte Primärsequenz der NOST-Ketolase aus *Nostoc sp. PCC 7120* kodiert

Die DNA, die für die NOST-Ketolase aus *Nostoc sp. PCC 7120* kodiert, wurde mittels PCR aus *Nostoc sp. PCC 7120* (Stamm der "Pasteur Culture Collection of Cyanobacterium") amplifiziert.

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc sp. PCC 7120*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO<sub>3</sub>, 0.04 g/l K<sub>2</sub>PO<sub>4</sub>·3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>·xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l EDTA disodium magnesium, 0.04 g/l Na<sub>2</sub>CO<sub>3</sub>, 1ml trace metal mix A5+Co (2.86 g/l H<sub>3</sub>BO<sub>3</sub>, 1.81 g/l MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.222 g/l ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.39 g/l NaMoO<sub>4</sub>·xH<sub>2</sub>O, 0.079 g/l CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.0494 g/l Co(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

Protokoll für DNA Isolation aus *Nostoc PCC7120*:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10minütige Zentrifugation bei 8 000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris HCl (pH 7.5) resuspendiert und in ein Eppendorf Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C

gelöst.

- Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc PCC 7120*, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc sp. PCC 7120* unter Verwendung eines
- 5 sense-spezifischen Primers (NOSTF, SEQ ID No. 79) und eines antisense-spezifischen Primers (NOSTG SEQ ID No. 80) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- 10 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:
- 1 µl einer *Nostoc sp. PCC 7120* DNA (hergestellt wie oben beschrieben)
  - 15 - 0.25 mM dNTPs
  - 0.2 mM NOSTF (SEQ ID No. 79)
  - 0.2 mM NOSTG (SEQ ID No. 80)
  - 5 µl 10X PCR-Puffer (TAKARA)
  - 0.25 µl R-Taq Polymerase (TAKARA)
  - 20 - 25.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten
- 25 35X94°C 1 Minute
- 55°C 1 Minuten
- 72°C 3 Minuten
- 1X72°C 10 Minuten

- 30 Die PCR-Amplifikation mit SEQ ID No. 79 und SEQ ID No. 80 resultierte in einem 805 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 81). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pGEM-T (Promega) kloniert und der Klon pNOSTF-G erhalten.

- 35 Sequenzierung des Klons pNOSTF-G mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 88,886-89,662 des Datenbankeintrages AP003592 identisch ist. Diese Nukleotidsequenz wurde in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im
- 40 verwendeten *Nostoc sp. PCC 7120*.

## Beispiel 2

Amplifikation einer DNA, die die gesamte Primärsequenz der NP196-Ketolase aus

5 *Nostoc punctiforme* ATCC 29133 kodiert

Die DNA, die für die NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 kodiert, wurde mittels PCR aus *Nostoc punctiforme* ATCC 29133 (Stamm der "American Type Culture Collection") amplifiziert.

10

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc punctiforme* ATCC 29133, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO<sub>3</sub>, 0.04 g/l K<sub>2</sub>PO<sub>4</sub>·3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>·xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l EDTA disodium magnesium, 0.04 g/l Na<sub>2</sub>CO<sub>3</sub>, 1ml Trace Metal Mix "A5+Co" (2.86 g/l H<sub>3</sub>BO<sub>3</sub>, 1.81 g/l MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.222 g/l ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.39 g/l Na-MoO<sub>4</sub>·2H<sub>2</sub>O, 0.079 g/l CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.0494 g/l Co(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

15

20

Protokoll für die DNA-Isolation aus *Nostoc punctiforme* ATCC 29133:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris-HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol-gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

30

35

Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme* ATCC 29133, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme* ATCC 29133 unter Verwendung eines sense-spezifischen Primers (NP196-1, SEQ ID No. 82) und

eines antisense-spezifischen Primers (NP196-2 SEQ ID No. 83) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- 5 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:
- 1 µl einer *Nostoc punctiforme* ATCC 29133 DNA (hergestellt wie oben beschrieben)
  - 10 - 0.25 mM dNTPs
  - 0.2 mM NP196-1 (SEQ ID No. 82)
  - 0.2 mM NP196-2 (SEQ ID No. 83)
  - 5 µl 10X PCR-Puffer (TAKARA)
  - 0.25 µl R Taq Polymerase (TAKARA)
  - 15 - 25.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten
- 20 35X94°C 1 Minute
- 55°C 1 Minuten
- 72°C 3 Minuten
- 1X72°C 10 Minuten

- 25 Die PCR-Amplifikation mit SEQ ID No. 82 und SEQ ID No. 83 resultierte in einem 792 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP196, SEQ ID No. 84). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP196 erhalten.

- 30 Sequenzierung des Klons pNP196 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 140.571-139.810 des Datenbankeintrages NZ\_AABC01000196 identisch ist (inverse orientiert zum veröffentlichten Datenbankeintrag) mit der Ausnahme, daß G in Position 140.571 durch A ersetzt wurde,
- 35 um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme* ATCC 29133.

- 40 Dieser Klon pNP196 wurde daher für die Klonierung in den Expressionsvektor pJIT117(Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

pJIT117 wurde modifiziert, indem der 35S-Terminator durch den OCS-Terminator (Octopine Synthase) des Ti-Plasmides pTi15955 von *Agrobacterium tumefaciens* (Datenbankeintrag X00493 von Position 12,541-12,350, Gielen et al. (1984) EMBO J. 3 835-846) ersetzt wurde.

Das DNA-Fragment, das die OCS-Terminatorregion beinhaltet, wurde mittels PCR unter Verwendung des Plasmides pHELLSGATE (Datenbankeintrag AJ311874, Wesley et al. (2001) Plant J. 27 581-590, nach Standardmethoden aus *E.coli* isoliert) sowie der Primer OCS-1 (SEQ ID No. 85) und OCS-2 (SEQ ID No. 86) hergestellt.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die die Octopin Synthase (OCS) Terminatorregion (SEQ ID No. 87) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten waren:

- 100 ng pHELLSGATE plasmid DNA
- 0.25 mM dNTPs
- 0.2 mM OCS-1 (SEQ ID No. 85)
- 0.2 mM OCS-2 (SEQ ID No. 86)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten
- 35X 94°C 1 Minute
- 50°C 1 Minute
- 72°C 1 Minute
- 1X72°C 10 Minuten

Das 210 bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pOCS erhalten.

Sequenzierung des Klons pOCS bestätigte eine Sequenz, die mit einem Sequenzabschnitt auf dem Ti-Plasmid pTi15955 von *Agrobacterium tumefaciens* (Datenbankeintrag X00493) von Position 12.541 bis 12.350 übereinstimmt.

Die Klonierung erfolgte durch Isolierung des 210 bp Sall-XhoI Fragmentes aus pOCS und Ligierung in den Sall-XhoI geschnittenen Vektor pJIT117.

Dieser Klon heisst pJO und wurde daher für die Klonierung in den Expressionsvektor pJONP196 verwendet.

5

Die Klonierung erfolgte durch Isolierung des 782 Bp SphI-Fragmentes aus pNP196 und Ligierung in den SphI geschnittenen Vektor pJO. Der Klon, der die NP196-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP196.

10

Beispiel 3:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Tagetes erecta*

15

Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus *Petunia hybrida* (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).

20

Das DNA Fragment, das die EPSPS Promoterregion (SEQ ID No. 88) aus *Petunia hybrida* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Petunia hybrida* isoliert) sowie der Primer EPSPS-1 (SEQ ID No. 89) und EPSPS-2 (SEQ ID No. 90) hergestellt.

25

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das EPSPS-Promoterfragment (Datenbankeintrag M37029: Nukleotidregion 7-1787) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

30

- 100 ng genomischer DNA aus *A.thaliana*
- 0.25 mM dNTPs
- 0.2 mM EPSPS-1 (SEQ ID No. 89)
- 0.2 mM EPSPS-2 (SEQ ID No. 90)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

35

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten  
35X94°C 1 Minute  
5 50°C 1 Minute  
72°C 2 Minute  
1X72°C 10 Minuten

- Das 1773 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pEPSPS erhalten.

- Sequenzierung des Klons pEPSPS bestätigte eine Sequenz, die sich lediglich durch zwei Deletion (Basen ctaagtttcagga in Position 46-58 der Sequenz M37029; Basen aaaaatat in Position 1422-1429 der Sequenz M37029) und die Basenaustausche (T statt G in Position 1447 der Sequenz M37029; A statt C in Position 1525 der Sequenz M37029; A statt G in Position 1627 der Sequenz M37029) von der publizierten EPSPS-Sequenz (Datenbankeintrag M37029: Nukleotidregion 7-1787) unterscheidet. Die zwei Deletionen und die zwei Basenaustausche an den Positionen 1447 und 1627 der Sequenz M37029 wurden in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentieren somit die tatsächliche Nukleotidsequenz in den verwendeten *Petunia hybrida* Pflanzen.

- Der Klon pEPSPS wurde daher für die Klonierung in den Expressionsvektor pJONP196 verwendet.

- Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJONP196. Der Klon, der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heisst pJOESP:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

- Zur Herstellung des Expressionsvektors MSP107 wurde das 2.961 KB bp SacI-XhoI Fragment aus pJOESP:NP196 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert. MSP 107 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment NP196 KETO CDS (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.



Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 5 Zur Herstellung des Expressionsvektors MSP108 wurde das 2.961 KB bp SacI-XhoI Fragment aus pJOESP:NP196 mit dem SacI-XhoI geschnittenen Vektor pSUN5 ligiert. MSP 108 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 10

Beispiel 4:

Amplifikation einer DNA, die die gesamte Primärsequenz der NP195-Ketolase aus *Nostoc punctiforme* ATCC 29133 kodiert

15

Die DNA, die für die NP195-Ketolase aus *Nostoc punctiforme* ATCC 29133 kodiert, wurde mittels PCR aus *Nostoc punctiforme* ATCC 29133 (Stamm der "American Type Culture Collection") amplifiziert. Die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc punctiforme* ATCC 29133 wurde in Beispiel 19 beschrieben.

20

Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme* ATCC 29133, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme* ATCC 29133 unter Verwendung eines sense-spezifischen Primers (NP195-1, SEQ ID No. 91) und eines antisense-spezifischen Primers (NP195-2 SEQ ID No. 92) amplifiziert.

25

Die PCR-Bedingungen waren die folgenden:

30

Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

35

- 1 µl einer *Nostoc punctiforme* ATCC 29133 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NP195-1 (SEQ ID No. 91)
- 0.2 mM NP195-2 (SEQ ID No. 92)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0.25 µl R Taq Polymerase (TAKARA)
- 25.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- 1X94°C 2 Minuten
- 35X94°C 1 Minute
- 5 55°C 1 Minuten
- 72°C 3 Minuten
- 1X72°C 10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 91 und SEQ ID No. 92 resultierte in einem 819 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP195, SEQ ID No. 93). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP195 erhalten.

Sequenzierung des Klons pNP195 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 55,604-56,392 des Datenbankeintrages NZ\_AABC010001965 identisch ist, mit der Ausnahme, daß T in Position 55.604 durch A ersetzt wurde, um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme* ATCC 29133.

Dieser Klon pNP195 wurde daher für die Klonierung in den Expressionsvektor pJ0 (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 809 Bp SphI-Fragmentes aus pNP195 und Ligierung in den SphI geschnittenen Vektor pJ0. Der Klon, der die NP195-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP195.

Beispiel 5:  
Amplifikation einer DNA, die die gesamte Primärsequenz der NODK-Ketolase aus *Nodularia spumignea* NSOR10 kodiert.

Die DNA, die für die Ketolase aus *Nodularia spumignea* NSOR10 kodiert, wurde mittels PCR aus *Nodularia spumignea* NSOR10 amplifiziert.

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nodularia spumignea* NSOR10, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in BG 11-Medium (1.5 g/l NaNO<sub>3</sub>, 0.04 g/l K<sub>2</sub>PO<sub>4</sub>·3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>·xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l

EDTA disodium magnesium, 0.04 g/l  $\text{Na}_2\text{CO}_3$ , 1ml Trace Metal Mix "A5+Co" (2.86 g/l  $\text{H}_3\text{BO}_3$ , 1.81 g/l  $\text{MnCl}_2 \times 4\text{H}_2\text{O}$ , 0.222 g/l  $\text{ZnSO}_4 \times 7\text{H}_2\text{O}$ , 0.39 g/l  $\text{NaMoO}_4 \times 2\text{H}_2\text{O}$ , 0.079 g/l  $\text{CuSO}_4 \times 5\text{H}_2\text{O}$ , 0.0494 g/l  $\text{Co}(\text{NO}_3)_2 \times 6\text{H}_2\text{O}$ ) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

5

Protokoll für die DNA-Isolation aus *Nodularia spumignea* NSOR10:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris\_HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

Die Nukleinsäure, kodierend eine Ketolase aus *Nodularia spumignea* NSOR10, wurde mittels "polymerase chain reaction" (PCR) aus *Nodularia spumignea* NSOR10 unter Verwendung eines sense-spezifischen Primers (NODK-1, SEQ ID No. 94) und eines antisense-spezifischen Primers (NODK-2 SEQ ID No. 95) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl einer *Nodularia spumignea* NSOR10 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NODK-1 (SEQ ID No. 94)
- 0.2 mM NODK-2 (SEQ ID No. 95)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0.25 µl R Taq Polymerase (TAKARA)
- 25.8 µl Aq. Dest.

40

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

35X94°C 1 Minute

5 55°C 1 Minuten

72°C 3 Minuten

1X72°C 10 Minuten

10 Die PCR-Amplifikation mit SEQ ID No. 94 und SEQ ID No. 95 resultierte in einem 720 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NODK, SEQ ID No. 96). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNODK erhalten.

15 Sequenzierung des Klons pNODK mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 2130-2819 des Datenbank-eintrages AY210783 identisch ist (inverse orientiert zum veröffentlichten Datenbankeintrag). Diese Nukleotidsequenz wurde in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nodularia spumigena* NSOR10.  
20

Dieser Klon pNODK wurde daher für die Klonierung in den Expressionsvektor pJ0 (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 710 Bp SphI-Fragmentes aus pNODK und Ligierung in den SphI geschnittenen Vektor pJ0.  
25 Der Klon, der die NODK-Ketolase von *Nodularia spumigena* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONODK.

30 Beispiel 6:  
Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NODK-Ketolase aus *Nodularia spumigena* NSOR10 in *Lycopersicon esculentum* und *Tagetes erecta*.

35 Die Expression der NODK-Ketolase aus *Nodularia spumigena* NSOR10 in *L. esculentum* und *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus *Petunia hybrida* (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).  
40

Der Klon pEPSPS (in Beispiel 8 beschrieben) wurde daher für die Klonierung in den Expressionsvektor pJONODK (in Beispiel 12 beschrieben) verwendet.

- Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJONODK. Der Klon, der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heisst pJOESP:NODK. Diese Expressionskassette enthält das Fragment NODK in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.
- 10 Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus *Nodularia spumignea* NSOR10 in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).
- 15 Zur Herstellung des Expressionsvektors MSP115 wurde das 2.889 KB bp SacI-XhoI Fragment aus pJOESP:NODK mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert. MSP 115 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea* NSOR10 NODK-Ketolase, 20 Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

- Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus *Nodularia spumignea* NSOR10 in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).
- 25

- Zur Herstellung des Expressionsvektors MSP116 wurde das 2.889 KB bp SacI-XhoI Fragment aus pJOESP:NODK mit dem SacI-XhoI geschnittenen Vektor pSUN5 ligiert. MSP 116 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea* NSOR10 NODK-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 30

35

Beispiel 6A:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Lycopersicon esculentum* und *Tagetes erecta*.

40

Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters PDS (Phytoendesaturase) aus *Lycopersicon esculentum* (Datenbankeintrag U46919).

5

Das DNA Fragment, das die PDS Promoterregion (SEQ ID No. 100) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer PDS-1 (SEQ ID No. 98) und PDS-2 (SEQ ID No. 99) hergestellt.

10

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das PDS-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

15

- 100 ng genomischer DNA aus *Lycopersicon esculentum*
- 0.25 mM dNTPs
- 0.2 mM PDS-1 (SEQ ID No. 98)
- 0.2 mM PDS-2 (SEQ ID No. 99)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

20

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

25

- |     |      |                 |
|-----|------|-----------------|
| 1X  | 94°C | 2 Minuten       |
| 35X | 94°C | 1 Minute        |
|     | 50°C | 1 Minute        |
|     | 72°C | 2 Minute        |
| 30  | 1X   | 72°C 10 Minuten |

30

Das 2096 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pPDS erhalten.

35

Der Klon pPDS wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 3 beschrieben) verwendet.

40

Die Klonierung erfolgte durch Isolierung des 2094 Bp Ecl136II-SmaI Fragmentes aus pPDS und Ligierung in den Ecl136II-HindIII geschnittenen Vektor pJOEPS:NP196. Die HindIII-Schnittstelle des Vektors wurde zuvor durch Behandlung mit dem Klenow-

Enzym in eine „blunt-end“-Schnittstelle überführt. Der Klon, der den Promoter PDS anstelle des ursprünglichen Promoters EPSPS enthält, heisst pJOPDS:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem *rbcS*-Transitpeptid.

5

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

- 10 Zur Herstellung des Expressionsvektors MSP117 wurde das 3.3 KB *Ecl*136II-*Xho*I Fragment aus pJOPDS:NP196 mit dem *Ecl*136II-*Xho*I geschnittenen Vektor pSUN3 ligiert. MSP 117 beinhaltet Fragment PDS den PDS Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

15

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

20

- 11 Zur Herstellung des Expressionsvektors MSP118 wurde das 3.3 KB bp *Ecl*136II-*Xho*I Fragment aus pJOPDS:NP196 mit dem *Ecl*136II-*Xho*I geschnittenen Vektor pSUN5 ligiert. MSP 118 beinhaltet Fragment PDS den PDS Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

25

Beispiel 6B:

- Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Lycopersicon esculentum* und *Tagetes erecta*.

30

- Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters B-GENE (chromoplastenspezifische lycopene B-cyclase) aus *Lycopersicon esculentum* (Datenbankeintrag AAZ51517).

35

- Das DNA Fragment, das die B-GENE Promoterregion (SEQ ID No. 103) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA

40

(nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer BGEN-1 (SEQ ID No. 101) und BGEN-2 (SEQ ID No. 102) hergestellt.

Die PCR-Bedingungen waren die folgenden:

5

Die PCR zur Amplifikation der DNA, die das B-GENE-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

☐

- 100 ng genomischer DNA aus *Lycopersicon esculentum*
- 10 - 0.25 mM dNTPs
- 0.2 mM BGEN-1 (SEQ ID No. 101)
- 0.2 mM BGEN-2 (SEQ ID No. 102)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 15 - 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- |    |      |                 |
|----|------|-----------------|
| 1X | 94°C | 2 Minuten       |
| 20 | 35X  | 94°C 1 Minute   |
|    |      | 50°C 1 Minute   |
|    |      | 72°C 2 Minute   |
|    | 1X   | 72°C 10 Minuten |

- 25 Das 1222 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pB-GENE erhalten.

Der Klon pB-GENE wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 3 beschrieben) verwendet.

30

Die Klonierung erfolgte durch Isolierung des 1222 Bp SacI-HindIII Fragmentes aus pB-GENE und Ligierung in den SacI-HindIII geschnittenen Vektor pJOEPS:NP196. Der Klon, der den Promoter B-GENE anstelle des ursprünglichen Promoters EPSPS enthält, heisst pJOBGEN:NP196. Diese Expressionskassette enthält das Fragment

35 NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

- Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).
- 40



Zur Herstellung des Expressionsvektors MSP119 wurde das 2.4 KB SacI-XhoI Fragment aus pJOBGEN:NP196 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert. MSP 119 beinhaltet Fragment B-GENE den B-GENE Promoter, Fragment *rbcS TP*  
5 *FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).  
10

Zur Herstellung des Expressionsvektors MSP120 wurde das 2.4 KB bp SacI-XhoI Fragment aus pJOBGEN:NP196 mit dem SacI-XhoI geschnittenen Vektor pSUN5 li-  
15 giert. MSP 120 beinhaltet Fragment B-GENE den B-GENE Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

20 Beispiel 6C:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus *Nostoc punctiforme* ATCC 29133 in *Lycopersicon esculentum* und *Tagetes erecta*.

25 Die Expression der NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters CHRC (chromoplast-specific carotenoid-associated protein) aus *Cucumis sativa* (Datenbank-eintrag AF099501).  
30

Das DNA Fragment, das die CHRC Promoterregion (SEQ ID No. 106) aus *Lycopersicon esculentum* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Lycopersicon esculentum* isoliert) sowie der Primer CHRC-1 (SEQ ID No. 104) und CHRC-2 (SEQ ID No. 105) hergestellt.  
35

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das CHRC-Promoterfragment beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:  
40

- 100 ng genomischer DNA aus *Lycopersicon esculentum*
- 0.25 mM dNTPs
- 0.2 mM CHRC-1 (SEQ ID No. 101)
- 0.2 mM CHRC-2 (SEQ ID No. 102)
- 5 - 5 µl 10X PCR-Puffer (Stratagene)
- 0.25 µl Pfu Polymerase (Stratagene)
- 28.8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

10

1X 94°C 2 Minuten

35X 94°C 1 Minute

50°C 1 Minute

72°C 2 Minute

15 1X 72°C 10 Minuten

Das 1222 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pCHRC erhalten.

- 20 Der Klon pB-GENE wurde daher für die Klonierung in den Expressionsvektor pJOEPS:NP196 (in Beispiel 2 beschrieben) verwendet.

Die Klonierung erfolgte durch Isolierung des 1540 Bp SacI-HindIII Fragmentes aus pCHRC und Ligierung in den SacI-HindIII geschnittenen Vektor pJOEPS:NP196. Der

- 25 Klon, der den Promoter CHRC anstelle des ursprünglichen Promoters EPSPS enthält, heisst pJOCHRC:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

- 30 Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

Zur Herstellung des Expressionsvektors MSP121 wurde das 2.6 KB SacI-XhoI Fragment aus pJOCHRC:NP196 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert.

- 35 MSP 121 beinhaltet Fragment CHRC den CHRC Promoter, Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der PDS-kontrollierten NP196-Ketolase aus *Nostoc punctiforme* in *Tagetes erecta* erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

- 5 Zur Herstellung des Expressionsvektors MSP122 wurde das 2.6 KB bp *SacI*-*XhoI* Fragment aus pJOCHRC:NP196 mit dem *SacI*-*XhoI* geschnittenen Vektor pSUN5 ligiert. MSP 122 beinhaltet Fragment CHRC den CHRC Promoter, Fragment *rbcS TP FRAGMENT* das *rbcS* Transitpeptid aus Erbse (194 bp), Fragment NP196 KETO CDS (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.
- 10

#### Beispiel 7:

#### Herstellung transgener *Tagetes* Pflanzen

- 15 *Tagetes* Samen werden sterilisiert und auf Keimungsmedium (MS-Medium; Murashige and Skoog, *Physiol. Plant.* 15(1962), 473-497) pH 5,8, 2% Saccharose) aufgelegt. Die Keimung erfolgt in einem Temperatur/Licht/Zeitintervall von 18-28°C/20-200 µE/3 - 16 Wochen, bevorzugt jedoch bei 21°C, 20-70 µE, für 4-8 Wochen.
- 20 Alle Blätter der sich bis dahin entwickelten *in vitro* Pflanzen werden geerntet und quer zur Mittelrippe geschnitten. Die dadurch entstehenden Blattexplantate mit einer Größe von 10 - 60 mm<sup>2</sup> werden im Verlaufe der Präparation in flüssigem MS - Medium bei Raumtemperatur für maximal 2 h aufbewahrt.
- 25 Ein beliebiger Agrobakterium *tumefaciens* Stamm, bevorzugt aber ein supervirulenter Stamm, wie z.B. EHA105 mit einem entsprechenden Binärplasmid, das ein Selektionsmarkergen (bevorzugt *bar* oder *pat*) sowie ein oder mehrere Trait- oder Reportergene tragen kann wird (pS5FNR:NOST, pS5AP3:NOST pS5FNR:NP196, pS5EPS:NP196, pS5FNR:NP195, pS5EPS:NP195, pS5FNR:NODK und
- 30 pS5EPS:NODK), über Nacht angezogen und für die Co-Kultivierung mit dem Blattmaterial verwendet. Die Anzucht des Bakterienstammes kann wie folgt erfolgen: Eine Einzelkolonie des entsprechenden Stammes wird in YEB (0,1 % Hefeextrakt, 0,5 % Rindfleischextrakt, 0,5 % Pepton, 0,5 % Saccharose, 0,5 % Magnesiumsulfat x 7 H<sub>2</sub>O) mit 25 mg/l Kanamycin angeimpft und bei 28°C für 16 bis 20 h angezogen. Anschließend
- 35 wird die Bakteriensuspension durch Zentrifugation bei 6000 g für 10 min geerntet und derart in flüssigem MS Medium resuspendiert, daß eine OD<sub>600</sub> von ca. 0,1 bis 0,8 entstand. Diese Suspension wird für die Co-Kultivierung mit dem Blattmaterial verwendet.

- Unmittelbar vor der Co-Kultivierung wird das MS-Medium, in dem die Blätter aufbewahrt worden sind, durch die Bakteriensuspension ersetzt. Die Inkubation der Blätt-
- 40

chen in der Agrobaktériensuspension erfolgte für 30 min unter leichtem Schütteln bei Raumtemperatur. Anschließend werden die infizierten Explantate auf ein mit Agar (z.B. 0,8 % Plant Agar (Duchefa, NL) verfestigtes MS-Medium mit Wachstumsregulatoren, wie beispielsweise 3 mg/l Benzylaminopurin (BAP) sowie 1 mg/l Indolyllessigsäure (IAA) aufgelegt. Die Orientierung der Blätter auf dem Medium ist bedeutungslos. Die Kultivierung der Explantate findet für 1 bis 8 Tage, bevorzugt aber für 6 Tage statt, dabei können folgende Bedingungen angewendet werden: Lichtintensität: 30 – 80  $\mu\text{Mol/m}^2 \times \text{sec}$ , Temperatur: 22 – 24°C, hell/dunkel Wechsel von 16/8 Stunden. Anschließend werden die co-kultivierten Explantate auf frisches MS-Medium, bevorzugt mit den gleichen Wachstumsregulatoren übertragen, wobei dieses zweite Medium zusätzlich ein Antibiotikum zur Unterdrückung des Bakterienwachstums enthält. Timentin in einer Konzentration von 200 bis 500 mg/l ist für diesen Zweck sehr geeignet. Als zweite selektive Komponente wird eine für die Selektion des Transformationserfolges eingesetzt. Phosphinothricin in einer Konzentration von 1 bis 5 mg/l selektiert sehr effizient, aber auch andere selektive Komponenten gemäß des zu verwendenden Verfahrens sind denkbar.

Nach jeweils ein bis drei Wochen erfolgt der Transfer der Explantate auf frisches Medium bis sich Sproßknospen und kleine Sprosse entwickeln, die dann auf das gleiche Basalmedium einschließlich Timentin und PPT oder alternative Komponenten mit Wachstumsregulatoren, nämlich z.B. 0,5 mg/l Indolylbuttersäure (IBA) und 0,5 mg/l Gibberillinsäure  $\text{GA}_3$ , zur Bewurzelung übertragen werden. Bewurzelte Sprosse können ins Gewächshaus überführt werden.

Zusätzlich zu der beschriebenen Methode sind folgende vorteilhafte Modifikationen möglich:

Bevor die Explantate mit den Bakterien infiziert werden, können sie für 1 bis 12 Tage, bevorzugt 3 - 4, auf das oben beschriebene Medium für die Co-Kultur vorinkubiert werden. Anschließend erfolgt die Infektion, Co-Kultur und selektive Regeneration wie oben beschrieben.

Der pH Wert für die Regeneration (normalerweise 5,8) kann auf pH 5,2 gesenkt werden. Dadurch wird die Kontrolle des Agrobakterienwachstums verbessert.

Die Zugabe von  $\text{AgNO}_3$  (3 - 10 mg/l) zum Regenerationsmedium verbessert den Zustand der Kultur einschließlich der Regeneration selbst.

Komponenten, die die Phenolbildung reduzieren und dem Fachmann bekannt sind, wie z.B. Zitronensäure, Ascorbinsäure, PVP u.v.a.m., wirken sich positiv auf die Kultur aus.

Für das gesamte Verfahren kann auch flüssiges Kulturmedium Verwendung finden. Die Kultur kann auch auf handelsüblichen Trägern, die auf dem flüssigen Medium positioniert werden inkubiert werden.

Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

Mit pS5FNR:NOST wurde beispielsweise erhalten: MSP102-1, MSP102-2, MSP102-3,

Mit pS5AP3:NOST wurde beispielsweise erhalten: MSP104-1, MSP104-2, MSP104-3

Mit pS5FNR:NP196 wurde erhalten: MSP106-1, MSP106-2, MSP106-3

Mit pS5EPS:NP196 wurde erhalten: MSP108-1, MSP108-2, MSP108-3

Mit pS5FNR:NP195 wurde erhalten: MSP110-1, MSP110-2, MSP110-3

Mit pS5EPS:NP195 wurde erhalten: MSP112-1, MSP112-2, MSP112-3

Mit pS5FNR:NODK wurde erhalten: MSP114-1, MSP114-2, MSP114-3

Mit pS5EPS:NODK wurde erhalten: MSP116-1, MSP116-2, MSP116-3

Mit pS3PDS:NP196 wurde erhalten: MSP117-1, MSP117-2, MSP117-3

Mit pS5PDS:NP196 wurde erhalten: MSP118-1, MSP118-2, MSP118-3

Mit pS3CHRC:NP196 wurde erhalten: MSP119-1, MSP119-2, MSP119-3

Mit pS5CHRC:NP196 wurde erhalten: MSP120-1, MSP120-2, MSP120-3

Mit pS3BGEN:NP196 wurde erhalten: MSP121-1, MSP121-2, MSP121-3

Mit pS5BGEN:NP196 wurde erhalten: MSP122-1, MSP122-2, MSP122-3

Beispiel 8:

Enzymatische Lipase-katalysierte Hydrolyse von Carotinoidestern aus Pflanzenmaterial und Identifizierung der Carotinoide

## Allgemeine Arbeitsvorschrift

- a) Gemörsertes Pflanzenmaterial (z.B. Petalenmaterial) (30-100 mg Frischgewicht) wird mit 100% Aceton (dreimal 500µl; jeweils etwa 15 Minuten schütteln) extrahiert.
- 5 Das Lösungsmittel wird evaporiert. Carotinoide werden anschließend in 495 µl Aceton aufgenommen, 4,95 ml Kaliumphosphatpuffer (100 mM, pH7.4) zugegeben und gut gemischt. Danach erfolgt die Zugabe von ca. 17 mg Bile-Salze (Sigma) und 149 µl einer NaCl/CaCl<sub>2</sub>-Lösung (3M NaCl und 75 mM CaCl<sub>2</sub>). Die Suspension wird für 30 Minuten bei 37°C inkubiert. Für die enzymatische Hydrolyse der Carotinoidester wird 595
- 10 µl einer Lipaselösung (50 mg/ml Lipase Typ7 von *Candida rugosa* (Sigma)) zugegeben und unter Schütteln bei 37°C inkubiert. Nach etwa 21 Stunden erfolgte nochmals eine Zugabe von 595 µl Lipase mit erneuter Inkubation von mindestens 5 Stunden bei 37°C. Anschließend werden etwa ca. 700 mg Na<sub>2</sub>SO<sub>4</sub> in der Lösung gelöst. Nach Zugabe von 1800 µl Petrolether werden die Carotinoide durch kräftig Mischen in die organische
- 15 Phase extrahiert. Dieses Ausschütteln wird solange wiederholt, bis die organische Phase farblos bleibt. Die Petroletherfraktionen werden vereinigt und der Petrolether evaporiert. Freie Carotinoide werden in 100-120 µl Aceton aufgenommen. Mittels HPLC und C30-reverse phase-Säule können freie Carotinoide aufgrund von Retentionszeit und UV-VIS-Spektren identifiziert werden.

20

Die verwendeten Bile-Salze oder Gallensäuresalze sind 1:1 Mischungen von Cholat und Desoxycholat.

- b) Arbeitsvorschrift für Aufarbeitung, wenn nur geringe Mengen an Carotinoidestern im
- 25 Pflanzenmaterial vorhanden sind

- Alternativ kann die Hydrolyse der Carotinoidester durch Lipase aus *Candida rugosa* nach Trennung mittels Dünnschichtchromatographie erreicht werden. Dazu werden 50-100mg Pflanzenmaterial dreimal mit etwa 750µl Aceton extrahiert. Der Lösungsmittel-
- 30 telextrakt wird im Vakuum einrotiert (erhöhte Temperaturen von 40-50°C sind tolerabel). Danach erfolgt Zugabe von 300µl Petrolether:Aceton (Verhältnis 5:1) und gute Durchmischung. Schwebstoffe werden durch Zentrifugation (1-2 Minuten) sedimentiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt. Das verbleibende Rest wird erneut mit 200µl Petrolether:Aceton (Verhältnis 5:1) extrahiert und Schwebstoffe
- 35 werden durch Zentrifugation entfernt. Die beiden Extrakte werden zusammengeführt (Volumen 500µl) und die Lösungsmittel evaporiert. Der Rückstand wird in 30µl Petrolether:Aceton (Verhältnis 5:1) resuspendiert und auf eine Dünnschichtplatte (Silica-Gel 60, Merck) aufgetragen. Falls mehr als eine Auftragung für präparativ-analytische Zwecke erforderlich ist, sollten mehrere Aliquots mit jeweils 50-100 mg Frischgewicht in der

beschriebenen Weise für die dünnschichtchromatographische Trennung aufbereitet werden.

Die Dünnschichtplatte wird in Petrolether:Aceton (Verhältnis 5:1) entwickelt. Carotinoidbanden können visuell aufgrund ihrer Farbe identifiziert werden. Einzelne Carotinoidbanden werden ausgekratzt und können für präparativ-analytische Zwecke gepoolt werden. Mit Aceton werden die Carotinoide vom Silica-Material eluiert; das Lösungsmittel wird im Vakuum evaporiert. Zur Hydrolyse der Carotinoidester wird der Rückstand in 495µl Aceton gelöst, 17mg Bile-Salze (Sigma), 4,95ml 0.1M Kaliumphosphatpuffer (pH 7,4) und 149µl (3M NaCl, 75mM CaCl<sub>2</sub>) zugegeben. Nach guter Durchmischung wird 30min bei 37°C äquiliibriert. Danach erfolgt die Zugabe von 595µl Lipase von *Candida rugosa* (Sigma, Stammlösung von 50mg/ml in 5mM CaCl<sub>2</sub>). Über Nacht erfolgt die Inkubation mit Lipase unter Schütteln bei 37°C. Nach etwa 21 Stunden wird nochmals die gleiche Menge an Lipase zugegeben; für mindestens 5 Stunden wird nochmals bei 37°C unter Schütteln inkubiert. Dann erfolgt die Zugabe von 700mg Na<sub>2</sub>SO<sub>4</sub> (wasserfrei); mit 1800µl Petrolether wird für ca. 1 Minute ausgeschüttelt und die Mischung bei 3500 Umdrehungen/Minute für 5 Minuten zentrifugiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt und das Ausschütteln so lange wiederholt, bis die obere Phase farblos ist. Die vereinigte Petrolether-Phase wird im Vakuum eingeeengt (Temperaturen von 40-50°C sind möglich). Der Rückstand wird in 120µl Aceton, eventuell mittels Ultraschall, gelöst. Die gelösten Carotinoide können mittels HPLC unter Verwendung einer C30-Säule getrennt und anhand von Referenzsubstanzen quantifiziert werden.

Beispiel 9:  
HPLC-Analyse freier Carotinoide

Die Analyse der nach der Arbeitsvorschriften in Beispiel 15 erhaltenen Proben erfolgt unter folgenden Bedingungen:

Folgende HPLC-Bedingungen wurden eingestellt.

Trennsäule: ProntoSil C30-Säule, 250 x 4,6 mm, (Bischoff, Leonberg, Germany)

Flussrate: 1.0 ml/min

Eluenten: Laufmittel A - 100% Methanol

Laufmittel B - 80% Methanol, 0.2% Ammoniumacetat

Laufmittel C - 100% t-Butyl-methylether

Detektion: 300-530 nm

Gradientenprofil:

Zeit	Flussrate	% Laufmittel A	% Laufmittel B	% Laufmittel C
------	-----------	----------------	----------------	----------------

1.00	1.0	95.0	5.0	0
12.00	1.0	95.0	5.0	0
12.10	1.0	80.0	5.0	15.0
22.00	1.0	76.0	5.0	19.0
22.10	1.0	66.5	5.0	28.5
38.00	1.0	15.0	5.0	80.0
45.00	1.0	95.0	5.0	0
46.0	1.0	95.0	5.0	0

Einige typische Retentionszeiten für erfindungsgemäß gebildete Carotinoide sind z.B.:  
Violaxanthin 11, 7 min, Astaxanthin 17,7 min, Adonixanthin 19 min, Adonirubin 19,9  
min, Zeaxanthin 21 min.



## Patentansprüche

1. Verwendung eines Promotors, ausgewählt aus der Gruppe
  - 5 A) EPSPS Promotor
  - B) B-Gene Promotor
  - C) PDS Promotor und
  - D) CHRC Promotor
- 10 zur Expression von Genen in Pflanzen der Gattung Tagetes, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
2. Verwendung nach Anspruch 1, dadurch gekennzeichnet, dass die Expression spe-  
15 zifisch in Blüten erfolgt.
3. Verwendung nach Anspruch 2, dadurch gekennzeichnet, dass die Expression spe-  
zifisch in Petalen erfolgt.
- 20 4. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der EPSPS Promotor gemäß Anspruch 1
  - A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder
  - 25 A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder 3 aufweist oder
  - A3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert oder
  - 30 A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3) enthält.
5. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass  
35 der B-Gene Promotor gemäß Anspruch 1
  - B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder

- 5 B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist oder
- 10 B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder
- B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3) enthält.

6. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der PDS Promotor gemäß Anspruch 1

- 15 C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder
- C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder
- 20 C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder
- C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3) enthält.

25 7. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der CHRC Promotor gemäß Anspruch 1

- 30 D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
- D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
- 35 D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
- D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3) enthält.

8. Genetisch veränderte Pflanze der Gattung *Tagetes*, wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7.
9. Genetisch veränderte Pflanze nach Anspruch 8, dadurch gekennzeichnet, dass die Regulation der Expression von Genen in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7 dadurch erreicht wird, dass man
- a) eine oder mehrere Promotoren gemäß einem der Ansprüche 1 bis 7 in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer endogenen Gene unter der Kontrolle der eingebrachten Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
  - b) ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
  - c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft eine oder mehrere, zu exprimierende Gene, in die Pflanze einbringt.
10. Genetisch veränderte Pflanze der Gattung *Tagetes*, enthaltend einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft ein zu exprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung *Tagetes*, die in Wildtyppflanzen der Gattung *Tagetes* von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
11. Genetisch veränderte Pflanze nach einem der Ansprüche 8 bis 10, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nuk-

leinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Kohlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von  
5 Vitaminen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Co-faktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen, wobei die Gene gegebenenfalls weitere Regulationselemente enthalten können.

10

12. Genetisch veränderte Pflanze nach Anspruch 11, dadurch gekennzeichnet, dass man als zu exprimierende Gene Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden verwendet.

15

13. Genetisch veränderte Pflanze nach Anspruch 12, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-  
20 Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

20

25

30

14. Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von genetisch veränderten Pflanzen der Gattung Tagetes gemäß einem der Ansprüche 8 bis 13.

35

15. Verfahren zur Herstellung von Carotinoiden durch Kultivierung von genetisch veränderten Pflanzen gemäß einem der Ansprüche 8 bis 13, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.
16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass man nach dem Kultivieren die genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den genetisch veränderten Pflanzen isoliert.
17. Verfahren nach Anspruch 16, dadurch gekennzeichnet, dass man die Blüten der genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den Petalen der genetisch veränderten Pflanzen isoliert.
18. Verfahren nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, dass die Carotinoide ausgewählt sind aus der Gruppe Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

## SEQUENCE LISTING

<110> SunGene GmbH & Co. KGaA

<120> Promotoren zur Expression von Genen in Tagetes

<130> PF 55341

<160> 106

<170> PatentIn version 3.1

<210> 1

<211> 174

<212> DNA

<213> Petunia hybrida

<220>

<221> promoter

<222> (1)..(174)

<223> EPSPS Promotor

<400> 1

gagaacacag ctggaat tttt ttacaaaggt agttggtgaa gctagtcagc gaatccatt 60

accttccact ctacctaacc cccttcacca acaacaaatt tctgtaattt aaaaactagc 120

caaaaaagaa ctctctttta caaagagcca aagactcaat ctttactttc aaga 174

<210> 2

<211> 1781

<212> DNA

<213> Petunia hybrida

<220>

<221> promoter

<222> (1) .. (1781)

<223> EPSPS Promotor

<400> 2

```
tagcgcaatc ttatgtggta caaatcttga ttagtcgggc taagtttcag gaaaaaatg      60
atgtggccct acaaatgggt ggaggatggg agatttggct ctatctagag ttatgtgggt      120
gttgaagcat ttggttactc tctgctgtgg tagttggcat atccacattg tctccttcca      180
cttttatgac aattacgtga aagttatggg ttgttttgtc tatttttgtc gaggcctttc      240
ttttccttcc aggttggtga agatgggtcca attcgattag aataatgttt tgagctttag      300
catattctct ctcgtttaca cgattatagt aataatgata taggatgaca gaagttgaca      360
cataaatttt ttattctctc catttacttt aatccaaatc tcacctacc taaacttctt      420
taatatgtat tcaatagtct atccgagtaa attgtaaatt taacaacat tgataatatt      480
gacacctact aacatatact agtaaagaga atattaacat ggcacatata atttgatgca      540
aatgagtat gatgaaattt aaacccaaaa tctcttgatt ttgacagtgt caccttgact      600
tgtaaactaa taagtcattg tttagtggca gaaagacaaa ctcattccacc aactgtatag      660
caataaaaaa tagaagaatc ttcctgaggg aaagttttgg aaaaattaag agtggctgag      720
atttaatttc aacaggaatt agttccactt aacttttagg ttacgatata gtgctaatta      780
aataacttaa ttgtattaga tatttcttgc acctaaaaaa tttaaaaact gaaaaaagg      840
agcaatcaaa ataaacaaaa ggacaaaata agtgaaagg      acagccacca accctggcgg      900
ctcactgttt gttggttaaa acgtagactt acacctacca aaatctacaa ctaaatgag      960
gcaataatac ttgcccmaa attaccaaga aaagaaaaag aaaggaatcc cttaatatta     1020
ctctcctcca tttcacaata aatatcctag ttgacttaa attagagttt aaaaaatgaa     1080
agacgacttt taaaacttgt aatctaaaat aaatcatagt taaatgtgtg gctataaatc     1140
attgtattaa cggtaaagtg gtaagtttaa aagttaattg ttttcaaata taaaattgta     1200
ctatcattct ttttggaatg gactaataag aaaactatga catccattat ggagcggagg     1260
```

gagtatctcc ttttaacaat aacctttgtc ccttcaattc aattatcagt atgcaaacat 1320  
taaaaattat tattgatggt aagtaccaca tcctccttaa tgatagaatc atcgtagaac 1380  
gcttttccag gcacacattc aaactagtta gaccaaaaaa tatgtaccac acatcgaata 1440  
gtccagactt ctttgtttga atagtcgact acattggata atggaacttc tcgaattaac 1500  
ttcgaattag tcgagcccca aataatatat acgtcgggtg gaaaactata aaatgtttga 1560  
caaaaatgtc aaattaatat atcaatctgc aacaaccttt tcaccttgag aacacagctg 1620  
gaatttttta caaaggtagt tgggtgaagct agtcagcgaa tccattacc ttccactcta 1680  
cctaaccccc ttcaccaaca acaaatttct gtaatttaaa aactagccaa aaaagaactc 1740  
tcttttacia agagccaaag actcaatctt tactttcaag a 1781

<210> 3

<211> 1760

<212> DNA

<213> *Petunia hybrida*

<220>

<221> promoter

<222> (1)..(1760)

<223> EPSPS Promotor

<400> 3

tagcgcaatc ttatgtggta caaatcttga ttagtcggga aaaaatgatg tggccctaca 60  
aatgggttga ggatgggaga tttggctcta tctagagtta tgtggttgtt gaagcatttg 120  
gttactctct gctgtggtag ttggcatatc cacattgtct ccttccactt ttatgacaat 180  
tacgtgaaag ttatgggttg ttttgtctat ttttgtcgag gcctttcttt tccttccagg 240  
ttgttgaaga tggccaatt cgattagaat aatgttttga gcttttagcat attctctctc 300  
gtttacacga ttatagtaat aatgatatag gatgacagaa gttgacacat aaatttttta 360  
ttctctccat ttactttaat ccaaattctc cctaccctaa acttctttta tatgtattca 420  
atagtctatc cgagtaaatt gtaaatttta caaccattga taatattgac acctactaac 480



atatactagt aaagagaata ttaacatggc acatataatt tgatgcaaaa tgagtatgat 540  
gaaattttaa cccaaaatct cttgattttg acagtgtcac cttgacttgt taactaataa 600  
gtcatgtttt agtggcagaa agacaaactc atccaccaac tgtatagcaa taaaaaatag 660  
aagaatcttc ctgaggcaaa gttttggaaa aattaagagt ggctgagatt taatttcaac 720  
aggaattagt tccacttaac ttttaggtta cgatacagtg ctaattaaat aacttaattg 780  
tattagatat ttcttgcacc taaaaaattt aaaaactgaa aaaaggtagc aatcaaaata 840  
aacaaaagga caaaataagt gaaaggtaca gccaccaacc ctggcggctc actgtttgtt 900  
ggttaaaacg tagacttaca cctacaaaaa tctacaacta aaatgaggca ataatacttt 960  
gccccaaatt accaagaaaa gaaaaagaaa ggaatccctt aatattactc tcctccattt 1020  
cacaataaat atcctagttt gacttaaatt agagttaa aaatgaaaga cgacttttaa 1080  
aacttgtaat ctaaaataaa tcatagttaa atgtgtggct ataaatcatt gtattaacgg 1140  
taaagtggta agttttaaag ttaattgttt tcaaataaa aattgtacta tcattctttt 1200  
tggaatggac taataagaaa actatgacat ccattatgga gcggaggagg tatctccttt 1260  
taacaataac ctttgtccct tcaattcaat tatcagtatg caaacattaa aaattattat 1320  
tgatgttaag taccacatca tccttaatga tagaatcatc gtagaacgct tttccaggca 1380  
cacattcaaa ctagttagac cagtaccaca catcgaatat tccagacttc tttgtttgaa 1440  
tagtcgacta cattggataa tggaacttct cgaattaact tcgaattagt cgagcccaaa 1500  
ataatatata cgtcgggtgg aaaactataa aatgtttgac aaaaatgtca aattaatata 1560  
tcaatctgca acaacctttt caccttgaga acacagctga aattttttac aaaggtagtt 1620  
ggtgaagcta gtcagcgaat ccattacct tccactctac ctaacccctc tcaccaacaa 1680  
caaatttctg taatttaaaa actagccaaa aaagaactct cttttacaaa gagccaaaga 1740  
ctcaatcttt actttcaaga 1760

&lt;210&gt; 4

&lt;211&gt; 1210

&lt;212&gt; DNA

<213> *Lycopersicon esculentum*

&lt;220&gt;

&lt;221&gt; promoter

&lt;222&gt; (1)..(1210)

&lt;223&gt; B-Gene Promotor

&lt;400&gt; 4

```
gaattctctg aaaaggagca ccatatTTgc cgcactgtgg tTcatatttc caagtacatt      60
tagatgaact atatcatcag attgaaaggT tattgtataa tcaatccagt ggattctcgt      120
tctggcacct ttagaagtac atgtgcggaa aagaatgata aggtttgtat tgttgttgac      180
aaagcctgtt gccttttctca tttgtaaatg ttctgaacga ctctaaatt actcttaagg      240
tgtaaggTct tccgtgcctg tttgtaaata taatgctgtg ccgtgactta ccttttgtac      300
catttgttca aatgtatggc ctgaacacca gggTtgtcaa aaatgtctca tgcccgTTTT      360
attggTctga aaatggcgtg atgccaaatt ctgccgctcc acagtgagca tttcgatcta      420
ctggaaattg accaacttat tttatcactt gataactaaa caaaatccta ttaactttaa      480
tcatacattg tatttatacc gaaaaattta tgcataactc attaaattac ctttttttagc      540
agtcaaattc taaatcagtt tctaatttat caaaatggct tttatagggt ccattttcca      600
ctaataTacc tgccgtccat gcactgacta caaaacaaat acctcactat gtttgTtagt      660
gcttggtaat ataaaacctt ttcttttatg agaaagttca ccgagaataa ttttctattt      720
gtggcataat agtatatagt gcagattgac aagaatttaa ttttgcagtt gggcacatga      780
acaattttcc tcaaagttgt agaaagtact tttcattttc ttgtcacoga aaattattta      840
taattgaaat taaaaccgaa tgagctgcaa gattcaagtc gaattttcaa aagaattgac      900
caagaaaaaa ttcaaaaata tccccaccc cctaccaaac acatcctaaa gtgaggTata      960
gactgggact gggattggga aaagggtaaa atgctttcac tagcttagca aagattccac     1020
tttgTtagct atctttcttt ctcatTTcct tttttctttt tctttttttt gttatataag     1080
ccaaagtagg taccctaaaag catcaatatt ttgtattgct tggTgattcc tctgtagtcc     1140
agtatttcat tttctacaag ttocacctcc ctccataatt aaccattatc aatcttatac     1200
attctctata                                     1210
```

&lt;210&gt; 5

<211> 1599

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1)..(1599)

<223> B-Gene Promotor

<400> 5

atctcattgt atagcttgtc ttttgtttca gtcgtcttag gcttgggtta gttggtgttg	60
ctgtttcata cttctatcaa ccttgtgtga gttcctttat aaaatatgac tgttggagga	120
agtaatttac ctttagttcg actacatcaa gatttgcac attctcgtcc aagaaatctt	180
agtttgaagc cttttggtct ggtatatttg tcaatctgag cttcgcaact ttctcatgac	240
aggggtttgt tgacatgcct gattgtgttc ttcctttact tgataattgc tgcttgttgc	300
ggaggcatca ctctaccttc ctgcagatca tgaattctct gaaaaggagc accatatttg	360
ccgcactgtg gttcatattt ccaattacat ttagatgaac tatatcatca ggagtgaag	420
gttattgtat aatcaatcca gtggattctc gttctggcac ctttagaagt acatgtgcgg	480
aaaagaatga taaggtttgt attgttgttg acaaggcctg ttgcctttct catttgtaaa	540
tgttctgaac gactcctaaa ttactcttaa agtgtaagggt cttccgtgcc tgtttgata	600
tataatgctg tgccgtgact taccttttgt accatttgtt caaatgtatg gcctggacac	660
tagggttgtc aaaaatgtct catgacttca cccttctttc ttgtcttggg gcccgtttta	720
ttggtctgag aacggcgtga tgccaaattc tgccgtcca cagtgagcat ttcgatctac	780
tggaaattga ccaacttatt ttatcacttg ataactagag tctgggttca aacaaaatcc	840
aataacttca atcatacatt gtatttatat tgaaaaaatt atgcacaact cagtaaatta	900
cctttttttg cagtcaaaaa ttctagatca gtttctaatt aatcaaaatg gcctttatag	960
ggtcccagtt ccattaatat acctgccgtc catgactga ttacaagaca aatacctcac	1020
tatgtttgtt agtgcttggg aatataaaac cttttctttt atgagaaagt tcaccgaaaa	1080
taattttcta tttgtggcat aactagtatc gaagtatata gtgcagattg acaagaattt	1140

```

aattttgcag ttgggcacat gaacaatttt cctcaaagtt gtagaaaata tttttcattt 1200
tcttgtcacc gaaaattatt tataattgaa attgaaaccg aatgagctgc aagactcgag 1260
tcgaatttca aaaaaattga ccaactaaat atgaaaaaat ccgaatatat cccccacccc 1320
ctaccaaaca catcctaaag tgaggatatag actgggactg ggattgggaa aagggtaaaa 1380
tgctttcact agcttagcaa agattccact ttgttagcta tctttctttc tcatttcctt 1440
ttttcttttt cttttttttg ttatataagc caaagtaggt acccaaaagc atcaatattt 1500
tgtattgctt ggtgattcct ctttactcca gtatttcatt ttctacaagt tccacctccc 1560
tcataatta accattatca atcttataca ttttctata 1599

```

<210> 6

<211> 1204

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1) .. (1204)

<223> B-Gene Promotor

<400> 6

```

tctgaaaagg agcaccatat ttgccgcact gtgggtcata tttccaagta catttagatg 60
aactatatca tcagattgaa aggttattgt ataatcaatc cagtggattc tcgttctggc 120
accttagaa gtacatgtgc ggaaaagaat gataaggttt gtattgttgt tgacaaagcc 180
tgttgccttt ctcatTTgta aatgttctga acgactccta aattactcct aagggtgaag 240
gtcttccgtg cctgtttgta aatataatgc tgtgccgtga cttacctttt gtaccatttg 300
ttcaaatgta tggcctggac accagggttg tcaaaaatgt ctcatgcccg ttttattggT 360
ctgagaatgg cgtgatgcca aattctgccg ctccacagtg agcatttcga tctactggaa 420
attgaccaac ttattttatc acttgataac taaacaaaat cctattaact ttaatcatac 480
attgtattta taccgaaaaa gttatgcata actcagtaaa ttaccttttt tagcagtcaa 540

```

attctagatc agttttctaatt ttatcaaaat ggcttttata ggggtcccagt tccactaata 600  
tacctgccgt ccatgcactg actacaagac aaatacctca ctatgtttgt tagtgcttgg 660  
taatataaaa ccttttcttt tatgagaaag ttcaccgaga ataattttct atttgtggca 720  
taactagtat atagtgcaga ttgacaagaa ttttaattttg cagttgggca catgaacaat 780  
tttcctcaaa gttgtagaaa gtacttttca ttttcttgtc accgaaaatt atttataatt 840  
gaaattaaaa ccgaatgagc tgcaagattc aagtcgaatt tcaaaagaat tgaccaagaa 900  
aaaattcaaa aatatcccc accccctacc aaacacatcc taaagtgagg tatagactgg 960  
gactgggatt gggaaaaggg taaaatgctt tcactagctt agcaaagatt ccactttgtt 1020  
agctatcttt ctttctcatt tccttttttc ttttctttt ttttgttata taagccaaag 1080  
taggtacca aaagcatcaa tattttgtat tgcttggtga ttcctctgta gtccagtatt 1140  
tcattttcta caagttccac ctccctccat aattaacctat tatcaatctt atacattctc 1200  
tata 1204

<210> 7

<211> 2078

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1)..(2078)

<223> PDS Promotor

<400> 7

tttgccagta ttacaacagc ttatatgttg agcaggtaaa agcttcaatg ccctattctt 60  
tctacagtta tcaatgttgc tcgtctaata tctgggtgtc ttctcgaaat gtcaattggc 120  
ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgtcaa 180  
atattattaat ttcatagtgt tcagtctcaa ttcttttagct ggttcctcat agtaaagttg 240  
tctaatatga aatgaaaatg ttctgtgtgt tgtactaata ccttttcatg gttgtctata 300

gaacgtcgat gaagagccaa acagaaacta ttttgggctg cgatttctga taccattgta 360  
tctgaatgct ggggtgggagc tcatcagaag ctttacaatg ggtcacatat atggagccgg 420  
tatgaggaat gctgggaatc agttgcgttt cgcgtgctag gacttttcct tcttgggtatt 480  
tctgcccaca gccagttga ttacgtgaac tccgtcagac ttggaaagga gagaagtacc 540  
caaagtgcgt ctttttagaa atacttttgt cacaaaatag cgggggtttac agctacagaa 600  
gatcatgcag aaggcgtcca gtttagtttt tgaaggttgt ttggagttaa tttatctaaa 660  
gtaaacttaa atcagctttt tgtttatgag ttcagtgaac tatatgttca aataagactt 720  
ccctttgtag atatgtgttt tttttgttgt tgagcacttt gtgtgcattg gataaacccc 780  
caacgtgtaa tagctaccat acaagagaag taactcgcac tgtccatgtc ttatgtggct 840  
cgactcagaa agcattcagg gggattgata accaccctcc aaaccaactg aaccattgtg 900  
aataaccacc cttcaaatca accgagtctt cgtgaaggac aaatatgtgg ttttatatac 960  
attaaatttt gtttttacat gcttcctctt acttcttttag ttttcttgac catatcttgc 1020  
gtttttccct tctgtaattg acacttttct tcaaaccatc cagcaatgtg gaagcttgac 1080  
gattttctt cagagtagaa attgaaaaga atcaactaaa aaggatagtc cttcgatttg 1140  
atttcgggt taaaaataaa ctaataagaa tgagagagcg aataatagaa tattttgaaa 1200  
ttttaagat attcaactat gttaaattgc gttataaatt tcttaaatta gtagcaccta 1260  
atagtttagt tctcaaaagt caaaactact acataatgtg ctcatTTTTc acattaaaaat 1320  
gcctacatga tgtaaaagta aaactcgtag cattctacgt gttttactca actcaaacat 1380  
cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt tctttttttt 1440  
ttttaaaaaa atattttttt ttatatcaat ccaaattggc tccaatttat cataaattag 1500  
gtagaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg ctatggtggg 1560  
acgtgtcaaa ttttggattg tagccaaaca tgagatttga tttaaagga attggccaaa 1620  
tcaccgaaag caggcatctt catcataaat tagtttgttt atttatacag aattatacgc 1680  
ttttactagt tatagcattc ggtatctttt tctgggtaac tgccaaacca ccacaaattt 1740  
caagtttcca ttaactctt caacttcaac ccaaccaaattt ttatttgctt aattgtgcag 1800  
aaccactccc tatatcttct aggtgctttc attcgttccg aggtaagaaa agatttttgt 1860  
ttctttgaat gctttatgcc actcgtttaa cttctgaggt ttgtggatct tttaggcgac 1920  
tttttttttt ttgtatgta aaatttgttt cataaatgct tctcaacata aatcttgaca 1980

aagagaagga attttaccaa gtatttaggt tcagaaatgg ataattttct tactgtgaaa 2040  
tatccttatg gcagggttta ctgttatttt tcagtaaa 2078

<210> 8

<211> 1342

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1) .. (1342)

<223> PDS Promotor

<400> 8

gttttttttg ttgttgagca ctttgtgtgc attggataaa cccccaacgt gtaatagcta 60  
ccatacaaga gaagtaactc gcactgtcca tgtcttatgt ggctcgactc agaaagcatt 120  
caggggggatt gataaccacc ctccaaacca actgaaccat tgtgaataac cacccttcaa 180  
atcaaccgag tcctcgtgaa ggacaaatat gtggttttat atacattaaa ttttgttttt 240  
acatgcttcc tcttacttct ttagttttct tgaccatata ttgcgttttt cccttctgta 300  
attgacactt ttcttcaaac catccagcaa tgtggaagct tgacgatttt ccttcagagt 360  
agaaattgaa aagaatcaac taaaaaggat agtccttcga ttgattttcc ggcttaaaaa 420  
taaactaata agaatgagag agcgaataat agaataattt gaaattttta agatattcaa 480  
ctatgtttaa ttgcgttata aatttcttaa attagtagca cctaatagtt tagttctcaa 540  
aagtcaaaac tactacataa tgtgctcatt ttccacatta aaatgcctac atgatgtaaa 600  
agtaaaaactc gtagcattct acgtgtttta ctcaactcaa acatcctgtt cattttaata 660  
aacgtacgat gagcttctct ctccaatttt cttttctttt ttttttttaa aaaaatattt 720  
ttttttatat caatccaaat gggctccaat ttatcataaa ttaggtagaa acttagatat 780  
taaagaaaga aaagggttta tctcgcaagt gtggctatgg tgggacgtgt caaattttgg 840  
attgtagcca aacatgagat ttgattttaa gggaattggc caaatcaccg aaagcaggca 900

tcttcatcat aaattagttt gtttatttat acagaattat acgcttttac tagttatagc 960  
attcgggtatc tttttctggg taactgccaa accaccacaa atttcaagtt tccattttaac 1020  
tcttcaactt caaccacaacc aaatttattt gcttaattgt gcagaaccac tccctatatc 1080  
ttctaggtgc tttcattcgt tccgaggtaa gaaaagattt ttgtttcttt gaatgcttta 1140  
tgccactcgt ttaacttctg aggtttgtgg atcttttagg cgactttttt tttttttgta 1200  
tgtaaaattt gtttcataaa tgcttctcaa cataaatctt gacaaagaga aggaatttta 1260  
ccaagtattt aggttcagaa atggataatt ttcttactgt gaaatctcct tatggcaggt 1320  
tttactgtta tttttcagta aa 1342

<210> 9

<211> 1008

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1)..(1008)

<223> PDS Promotor

<400> 9  
aagcttgacg attttccttc agagtagaaa ttgaaaagaa tcaactaaaa aggatagtcc 60  
ttcgatttga tttccggctt aaaaataaac taataagaat gagagagcga ataatagaat 120  
attttgaaat tttaaagata ttcaactatg ttaaattgcg ttataaattt cttaaattag 180  
tagcacctaa tagtttagtt ctcaaaagtc aaactacta cataatcgtg ctcatTTTTT 240  
acattaaaaat gcttacatga tgtaaaagta aaactcgtag cattctacgt gttttactca 300  
actcaaacat cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt 360  
tctttttttt ttttaaaaaa atattttttt ttatatcaat ccaaattggc tccaatttat 420  
cataaattag gtagaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg 480  
ctatgggtggg acgtgtcaaa ttttggtattg tagccaaaca tgagatttga tttaaaggga 540



attggccaaa tcaccgaaag caggcatctt catcataaat tagtttggtt atttatacag 600  
aattatacgc ttttactagt tatagcattc ggtatctttt tctgggtaac tgccaaacca 660  
ccacaaaattt caagtttcca ttttaactctt caacttcaac ccaaccaaatt ttatttgctt 720  
aattgtgcag aaccactccc tatactcttct aggtgctttc attcgttccg aggtaagaaa 780  
agatttttgt ttctttgaat gctttatgcc actcgtttaa cttctgaggt ttgtggatct 840  
tttaggcgac tttttttttt ttgtatgta aaatttggtt cataaatgct tctcaacata 900  
aatcttgaca aagagaagga attttaccaa gtatttaggt tcagaaatgg ataattttct 960  
tactgtgaaa tatccttatg gcaggtttta ctgttatttt tcagtaaa 1008

<210> 10

<211> 2078

<212> DNA

<213> *Lycopersicon esculentum*

<220>

<221> promoter

<222> (1)..(2078)

<223>

<400> 10

tttgccagta ttacaacagc ttatatgttg agcaggtaaa agcttcaatg ccctattctt 60  
tctacagtta tcaatgttgc tcgtctaata tctgggtgtc ttctcgaaat gtcaattggc 120  
ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgtaa 180  
at ttattaat ttcatagtgt tcagtctcaa ttctttagct gtttctcat agtaaagttg 240  
tctaatatga aatgaaaatg ttctgtgtgt tgtactaata ccttttcatg gttgtctata 300  
gaacgtcgat gaagagccaa acagaaacta ttttgggctg cgatttctga taccattgta 360  
tctgaatgct ggggtgggagc tcatcagaag ctttacaatg ggtcacatat atggagccga 420  
gtatgaggaa tgctgggaat cagttgtgct tcgcgtgcta ggacttttcc ttctgggtat 480  
ttctgcccac agcccagttg attacgtgaa ctccgtcaga cttggaaagg agagaagtac 540

ccaaatgtcg tctttttaga aatacttttg tcacaaaata gcgggggttta cagctacaga 600  
agatcatgca gaaggcgctc agtttagttt ttgaaggttg tttggagttt atttatctaa 660  
agtaaaactta aatcagcttt ttgtttatga gttcagtga ctatatgttc aaataagact 720  
tccctttgta gaatatgtgt ttttttttgt tgttgagcac tttgtgtgca ttggataaac 780  
ccccaacgtg taatagctac catacaagag aagtaactcg cactgtccat gtcttatgtg 840  
gctcgactca gaaagcattc aggggggattg ataaccaccc tccaaaccaa ctgaaccatt 900  
gtgaataacc acccttcaaa tcaaccgagt cctcgtgaag gacaaatatg tggttttata 960  
tacattaaat tttgttttta catgcttcct cttacttctt tagttttctt gaccatatct 1020  
tctttttccc ttctgtaatt gacattttct tcaaaccatc cagcaatgtg gaagcttyac 1080  
gattttcctt cagagtagaa attgaaaaga atcaactaaa aaggatagtc cttcgatttg 1140  
atttccggct taaaaataaa ctaataagaa tgagagagcg aataatagaa tattttgaaa 1200  
ttttaaagat attcaactat gttaaattgc gttataaatt tcttaaatta gtagcaccta 1260  
atagtttagt tctcaaaagt caaaactact acataatgtg ctcatTTTTT acattaaaat 1320  
gcctacatga tgtaaaagta aaactcgtag cattctacgt gttttactca actcaaacat 1380  
cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt tctttttttt 1440  
ttttaaaaaa atattttttt ttatatcaat ccaaattggc tccaatttat cataaattag 1500  
gtagaaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg ctatgggtggg 1560  
acgtgtcaaa ttttggattg tagccaaaca tgagatttga tttaaaggga attggccaaa 1620  
tcaccgaaag caggcatctt catcataaat tagtttgttt atttatacag aattatacgc 1680  
ttttactagt tatagcattc ggtatctttt tctgggtaac tgccaaacca ccacaaattt 1740  
caagtttcca ttttaactctt caacttcaac ccaaccaaatttatttgcctt aattgtgcag 1800  
aaccactccc tatatcttct aggtgctttc attcgttccg aggtaagaaa agatttttgt 1860  
ttctttgaat gctttatgcc actcgtttta cttctgaggt ttgtggatct tttaggcgac 1920  
tttttttttt tttgtatgta aaatttgttt cataaatgct tctcaacata aatcttgaca 1980  
aagagaagga attttaccaa gtatttaggt tcagaaatgg ataattttct tactgtgaaa 2040  
tatccttatg gcagggttta ctgttatttt tcagtaaa 2078

&lt;211&gt; 1528

&lt;212&gt; DNA

&lt;213&gt; Cucumis sativus

&lt;220&gt;

&lt;221&gt; promoter

&lt;222&gt; (1)..(1528)

&lt;223&gt; CHRC Promotor

&lt;400&gt; 11

```
tacaaattag ggttacttta ttcattttca tccattctct ttattgtaa attttgta 60
tttattcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca 120
ctcaagcctt taccatcttc cttttctatt tcaatactat ttctacttca tttttcacgt 180
ttttaacatc tttctttatt tcttgctcac ttcgtttagg gatgcctaata gtcctaaatt 240
tcctctctcg tagtaacaca aaaccaatgt aatgctactt ctctctacat ttttaataca 300
aataaagtga aacaaaatat ctataaataa acaaataatat atattttggt agacgctgtc 360
tcaaccatc aattaaaaaa ttttggtata tttctacttt acctactaaa tttgtttctc 420
atattttacct ttttaaccccc acaaaaaaaaa attataaaaa agaaagaaaa aagctaaacc 480
ctattttaaat agctaactat aagatcttaa aattatcctc atcagtgtat agtttaattg 540
gttattaact tataacatta tatatctatg acatatactc tctcctagct atttctcaca 600
ttttttaact taagaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta 660
atttgattaa caaaaagtta gaaatattta tttaaataaa aaagactaat aaatatataa 720
aatgaatggt catagcgaga cccatttaga gatgagtatg ctttcacatg ctgagattat 780
tttcaaaaact aagggtgtag caatattaaa tcaataaaaat tattataaat aacaaaatta 840
acctgctcgt gtttgctgta tatgggaggg tacaaaataa attaaactaa agatgattat 900
gttttagaca ttttttctat ctgtattagt ttatacatat taattcagga gctgcacaac 960
ccaattctat tttcgttctt tgggtgctgg gtttctcaca aggttcaata gtcaatatta 1020
ggttttattg gacttttaaat agtatcaaac aaatctatgt gtgaacttaa aaattgtatt 1080
aaatatttag ggtaacctgt tgccgttttt agaataatgt ttcttcttaa tacacgaaag 1140
```

cgtattgtgt attcattcat ttggcgccctc acatgcttcg gttggctcgc tttagtctct 1200  
gccttctttg tatattgtac tccccctctt cctatgccac gtgttctgag cttacaaga 1260  
cacgttgctg gccattgcca aacaagtcac tttaacttca caaggccga tttgacctcc 1320  
aaaacaacga caagtttccg aacagtcgag aagatcaagg gtataatcgt ctttttgaat 1380  
tctatttctc tttatttaat agtccctctc gtgtgatagt ttttaaaaga tttttaaac 1440  
gtagctgctg tttaagtaaa tcccagtcct tcagtttctg cttttgtgtg ttttgttct 1500  
ctgatttacg gaatttggaa ataattct 1528

<210> 12

<211> 1538

<212> DNA

<213> *Cucumis sativus*

<220>

<221> promoter

<222> (1) .. (1538)

<223> CHRC Promotor

<400> 12

tacaaattag ggttacttta ttcatTTTca tccattctct ttattgTtaa attttgtaca 60  
tttattcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca 120  
ctcaagcctt taccatcttc cttttctatt tcaatactat ttctacttca tttttcacgt 180  
ttttaacatc tttctttatt tcttgtccac ttcgtttagg gatgcctaatt gtcccaaatt 240  
tcatctctcg tagtaacaca aaaccaatgt aatgctactt ctctctacat ttttaataca 300  
aataaagtga aacaaaatat ctataaataa acaaatatat atattttgtt agacgctgtc 360  
tcaaccocatc aattaaaaaa ttttgTTata tttctacttt acctactaaa tttgtttctc 420  
atattttacct tttaaccccc acaaaaaaaaa attataaaaa agaaagaaaa aagctaaacc 480  
ctattttaaat agctaactat aagatcttaa aattatcttc atcagtgtat agtttaattg 540  
gttattaact tataacatta tatatctatg acatatactc tctcctagct atttctcaca 600

ttttttaact taagaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta 660  
atttgattaa caaaaagtta gaaatattta tttaaataaa aaagactaat aaatatataa 720  
aatgaatgtt catacgcaga cccatttaga gatgagtatg ctttcacatg ctgagattat 780  
tttcaaaact aagggtgtag caatattaaa tcaataaaat tattataaat aacaaaatta 840  
acctgctcgt gtttgctgta tatgggaggc tacaaaataa attaaactaa agatgattat 900  
gttttagaca ttttttctat ctgtattagt ttatacatat taattcagga gctgcacaac 960  
ccaattctat tttcgttctt tgggtggctgg gtttctcaca aggttcaata gtcaatatta 1020  
ggttttattg gacttttaat agtatcaaac aaatctatgt gtgaacttaa aaattgtatt 1080  
aaatatttag ggtaacctgt tgccgttttt agaataatgt ttcttcttaa tacacgaaag 1140  
cgtattgtgt attcattcat ttggcgcttc acatgcttcg gttggctcgc tttagtctct 1200  
gccttctttg tattttgtac tccccctctt cctatgccac gtgttctgag cttacaaga 1260  
cacgttgctg gccattgcca aacaagtcac ttttaacttca caaggtccga tttgacctcc 1320  
aaaacaacga caagtttccg aacagtcgag aagatcaagg gtataatcgt ctttttgaat 1380  
tctatttctc tttatttaat agtccctctc gtgtgatagt tttttaaaga tttttaaaac 1440  
gtagctgctg ttttaagtaaa tcccagtcct tcagtttgtg cttttgtgtg ttttgtttct 1500  
ctgatttacg gaatttggaa ataattcttc taaagggc 1538

<210> 13

<211> 1525

<212> DNA

<213> *Cucumis sativus*

<220>

<221> promoter

<222> (1)..(1525)

<223> CHRC Promotor

<400> 13

tacaaattag ggttacttta ttcattttca tccattctct ttattgttaa attttgtaca 60

ttatttcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca 120  
 ctcaagcctt taccatcttc cttttctatt tcaatactat ttctacttca tttttcacgt 180  
 ttttaacatc tttctttatt tcttgccac ttcgtttagg gatgcctaata gtcctaaatt 240  
 tcatctctcg tagtaacaca aaaccaatgt aatgctacct ctctctacat ttttaataca 300  
 aataaagtga aacaaaatat ctataaataa acaaatatat atattttgtt agacgctgtc 360  
 tcaaccatc aattaataaa ttttggtata tttctacttt acctactaaa tttgtttctc 420  
 atattttacct tttaaccccc acaaaaaaaaa attataaaaa agaaagaaaa aagctaaacc 480  
 ctattttaaat agctagctat aagatcttaa aattatcttc atcagtgtat agtttaattg 540  
 gttatttaact tataacatta tataatctatg acatatactc tctcctagct atttctcaca 600  
 ttttttaact taagaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta 660  
 atttgattaa caaaaagata gaaatattta tttaaataaa aaagactaat aaatatataa 720  
 aatgaatgtt catacgcaga cccattttaga gatgagtatg ctttcacatg ttgagattat 780  
 tttcaaaact aaggttgtag caatattaaa tcaataaaat tattataaat aacaaaatta 840  
 acctgctcgt gtttgctgaa tatgggaggc taaaaataa attaaactaa agatgattat 900  
 gttttagaca ttttttctat ctgtattagt ttatacatat taattcagga gctgcacaac 960  
 ccaattctat tttcgttcct tgggtggctgg gtttctcaca aggttcaata gtcaatatta 1020  
 ggttttattg gacttttaaat agtatcaaata aaatctatgt gtgaacttaa aaattgtatt 1080  
 aaatatttag ggtaacctgt tgccgttttt agaataatgt ttcttcttaa tacacgaaag 1140  
 cgtattgtgt attcattcat ttggcgcttc acatgcttcg gttggctcgc tttagtctct 1200  
 gccttctttg tattttgtac tccccctctt cctatgccac gtgttctgag cttacaagc 1260  
 cacgttgctg gccattgcca aacaagtcac tttaacttca caaggtccga tttgacctcc 1320  
 aaaacaacga caagtttccg aacagtcgag aagatcaagg gtataatcgt ctttttgaat 1380  
 tctatttctc tttatttaaat agtcctctc gtgtgatagt tttttaaaga tttttaaac 1440  
 gtagctgctg ttttaagtaaa tcccagtcct tcagtttggtg cttttgtgtg tttgtttct 1500  
 ctgatttacg gaatttggaataaat 1525

&lt;210&gt; 14

&lt;211&gt; 1519

&lt;212&gt; DNA

&lt;213&gt; Cucumis sativus

&lt;220&gt;

&lt;221&gt; promoter

&lt;222&gt; (1)..(1519)

&lt;223&gt; CHRC Promotor

&lt;400&gt; 14

ttactttatt cattttcatc cattctcttt attgttaaatt ttgttacatt tattcaataa	60
tattatatgt ttattacaaa ttctcacttt ctatttcata cctattcact caagccttta	120
ccatcttccct tttctatttc aatactattt ctacttcatt ttccacgttt ttaacatctt	180
tctttatttc ttgtccactt cgtttaggga tgcctaattgt cccaaatttc atctctcgta	240
gtaacacaaa accaatgtaa tgctacctct ctctacattt ttaatacaaa taaagtga	300
caaaatatct ataaataaac aaatatatat attttgtagt acgctgtctc aacccatcaa	360
ttaaaaaatt ttgttatatt tctactttac ctactaaatt tgtttctcat atttaccttt	420
taacccccac aaaaaaaaaat tataaaaaag aaagaaaaaa gctaaaccct atttaa	480
ctagctataa gatcttaaaa ttatctctcat cagtgtatag tttaattggg tattaactta	540
taacattata tatctatgac atatactctc tcctagctat ttctcacatt ttttaactta	600
agaaaatagt cataacatag tctaaaattc aaacatccac atgctcta	660
tgattaaca	
aaaagataga aatatttatt taaataaaaa agactaataa atatataaaa tgaatgttca	720
tacgcagacc catttagaga tgagtatgct ttcacatggt gagattattt tcaaaactaa	780
ggttgtagca atattaaatc aataaaatta ttataaataa caaaattaac ctgctcgtgt	840
ttgctgaata tgggaggcta caaaataaat taaactaaag atgattatgt tttagacatt	900
ttttctatct gtattagttt atacatatta attcaggagc tgcacaaccc aattctattt	960
tcgttccttg gtggctgtgt ttctcacaag gttcaatagt caatattagg ttttattgga	1020
cttttaatag tatcaaataa atctatgtgt gaacttaaaa attgtattaa atatttaggg	1080
taacctgttg ccgttttttag aataatgttt cttcttaata cacgaaagcg tattgtgtat	1140
tcattcattt ggccctcac atgcttcggg tggctcgctt tagtctctgc cttctttgta	1200

```

ttttgtactc cccctcttcc tatgccacgt gttctgagct taacaagcca cgttgcgtgc 1260
cattgccaaa caagtcattt taacttcaca aggtccgatt tgacctccaa aacaacgaca 1320
agtttccgaa cagtcgcgaa gatcaagggt gtaatcgtct ttttgaattc tatttctctt 1380
tatttaatag tccctctcgt gtgatagttt tttaaagatt tttaaaacgt agctgctggt 1440
taagtaaadc ccagtccttc agtttgtgct tttgtgtggt ttgtttctct gatttacgga 1500
atttgaaaat aataagctt 1519

```

<210> 15

<211> 1771

<212> DNA

<213> *Haematococcus pluvialis*

<220>

<221> CDS

<222> (166)..(1155)

<223>

<400> 15

```

ggcacgagct tgcacgcaag tcagcgcgcg caagtcaaca cctgccgggc cacagcctca 60
aataataaag agctcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac 120
ccgcgagtct cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca 177
                                     Met Gln Leu Ala
                                     1

```

```

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
5          10          15          20

```

```

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
          25          30          35

```

```

gcg acc cag tac tgc ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro
          40          45          50

```

```

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369

```



Gly	Leu	Lys	Asn	Ala	Tyr	Lys	Pro	Pro	Pro	Ser	Asp	Thr	Lys	Gly	Ile		
	55						60					65					
aca	atg	gcg	cta	cgt	gtc	atc	ggc	tcc	tgg	gcc	gca	gtg	ttc	ctc	cac		417
Thr	Met	Ala	Leu	Arg	Val	Ile	Gly	Ser	Trp	Ala	Ala	Val	Phe	Leu	His		
	70					75					80						
gcc	att	ttt	caa	atc	aag	ctt	ccg	acc	tcc	ttg	gac	cag	ctg	cac	tgg		465
Ala	Ile	Phe	Gln	Ile	Lys	Leu	Pro	Thr	Ser	Leu	Asp	Gln	Leu	His	Trp		
	85				90					95					100		
ctg	ccc	gtg	tca	gat	gcc	aca	gct	cag	ctg	gtt	agc	ggc	acg	agc	agc		513
Leu	Pro	Val	Ser	Asp	Ala	Thr	Ala	Gln	Leu	Val	Ser	Gly	Thr	Ser	Ser		
				105					110					115			
ctg	ctc	gac	atc	gtc	gta	gta	ttc	ttt	gtc	ctg	gag	ttc	ctg	tac	aca		561
Leu	Leu	Asp	Ile	Val	Val	Val	Phe	Phe	Val	Leu	Glu	Phe	Leu	Tyr	Thr		
			120					125					130				
ggc	ctt	ttt	atc	acc	acg	cat	gat	gct	atg	cat	ggc	acc	atc	gcc	atg		609
Gly	Leu	Phe	Ile	Thr	Thr	His	Asp	Ala	Met	His	Gly	Thr	Ile	Ala	Met		
	135						140					145					
aga	aac	agg	cag	ctt	aat	gac	ttc	ttg	ggc	aga	gta	tgc	atc	tcc	ttg		657
Arg	Asn	Arg	Gln	Leu	Asn	Asp	Phe	Leu	Gly	Arg	Val	Cys	Ile	Ser	Leu		
	150					155					160						
tac	gcc	tgg	ttt	gat	tac	aac	atg	ctg	cac	cgc	aag	cat	tgg	gag	cac		705
Tyr	Ala	Trp	Phe	Asp	Tyr	Asn	Met	Leu	His	Arg	Lys	His	Trp	Glu	His		
	165				170					175					180		
cac	aac	cac	act	ggc	gag	gtg	ggc	aag	gac	cct	gac	ttc	cac	agg	gga		753
His	Asn	His	Thr	Gly	Glu	Val	Gly	Lys	Asp	Pro	Asp	Phe	His	Arg	Gly		
				185					190					195			
aac	cct	ggc	att	gtg	ccc	tgg	ttt	gcc	agc	ttc	atg	tcc	agc	tac	atg		801
Asn	Pro	Gly	Ile	Val	Pro	Trp	Phe	Ala	Ser	Phe	Met	Ser	Ser	Tyr	Met		
			200					205					210				
tcg	atg	tgg	cag	ttt	gcg	cgc	ctc	gca	tgg	tgg	acg	gtg	gtc	atg	cag		849
Ser	Met	Trp	Gln	Phe	Ala	Arg	Leu	Ala	Trp	Trp	Thr	Val	Val	Met	Gln		
		215					220					225					
ctg	ctg	ggt	gcg	cca	atg	gcg	aac	ctg	ctg	gtg	ttc	atg	gcg	gcc	gcg		897
Leu	Leu	Gly	Ala	Pro	Met	Ala	Asn	Leu	Leu	Val	Phe	Met	Ala	Ala	Ala		
		230				235					240						
ccc	atc	ctg	tcc	gcc	ttc	cgc	ttg	ttc	tac	ttt	ggc	acg	tac	atg	ccc		945
Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	Phe	Tyr	Phe	Gly	Thr	Tyr	Met	Pro		
	245				250					255					260		
cac	aag	cct	gag	cct	ggc	gcc	gcg	tca	ggc	tct	tca	cca	gcc	gtc	atg		993
His	Lys	Pro	Glu	Pro	Gly	Ala	Ala	Ser	Gly	Ser	Ser	Pro	Ala	Val	Met		
				265				270					275				
aac	tgg	tgg	aag	tcg	cgc	act	agc	cag	gcg	tcc	gac	ctg	gtc	agc	ttt		1041

Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe  
 280 285 290  
 ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc 1089  
 Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro  
 295 300 305  
 ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga 1137  
 Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg  
 310 315 320  
 ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcca 1185  
 Gly Leu Val Pro Ala  
 325  
 gctgggcatg cagggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg 1245  
 gacacgctgc atgggctacc ctgtgtagct gcgcgcaacta ggggaggggg tttgtagctg 1305  
 tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gagtacacc acaggccaac 1365  
 acccttgcag gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425  
 tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgccctgc 1485  
 aagggtgcagg cacaagctag gctggacgag gactcgggtg caggcagggtg aagaggtgcg 1545  
 ggaggggtggt gccacaccca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605  
 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665  
 agatactggt caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725  
 ccctgcgctt atgaagctgt aacaataaag tggttcaaaa aaaaaa 1771

&lt;210&gt; 16

&lt;211&gt; 329

&lt;212&gt; PRT

&lt;213&gt; Haematococcus pluvialis

&lt;400&gt; 16

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala  
 1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val  
 20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp  
 35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp  
 50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala  
 65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp  
 85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser  
 100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu  
 115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly  
 130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val  
 145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys  
 165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp  
 180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met  
 195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr  
 210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe  
 225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly  
 245 250 255

## 23

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser  
 260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp  
 275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His  
 290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg  
 305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala  
 325

<210> 17

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168)..(1130)

<223>

<400> 17

cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60

gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg 120

ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc 176  
 Met His Val  
 1

gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224  
 Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser  
 5 10 15

agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272  
 Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser  
 20 25 30 35

gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro 40 45 50	320
cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly 55 60 65	368
acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro 70 75 80	416
aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala 85 90 95	464
cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe 100 105 110 115	512
att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp 120 125 130	560
gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu 135 140 145	608
ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met 150 155 160	656
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly 165 170 175	704
aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe 180 185 190 195	752
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu 200 205 210	800
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn 215 220 225	848
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu 230 235 240	896
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala 245 250 255	944

gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca 992  
 Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala  
 260 265 270 275  
  
 tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg 1040  
 Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp  
 280 285 290  
  
 gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc 1088  
 Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys  
 295 300 305  
  
 cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga 1130  
 Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala  
 310 315 320  
  
 cctgggtccct ccgctggtga cccagcgtct gcacaagagt gtcagtgtac aggggtgctgc 1190  
  
 ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca 1250  
  
 ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310  
  
 ggcgtgtac tgacaatggg cgtgctactg gggctctggca gtggctagga tggagtttga 1370  
  
 tgcattcagt agcgggtggc aacgtcatgt ggatggtgga agtgctgagg ggtttaggca 1430  
  
 gccggcattt gagaggggcta agttataaat cgcagtgtgc tcatggcac atatctgcac 1490  
  
 acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggtt tcgtgctatt 1550  
  
 gttttattca gcagcagtac ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610  
  
 gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

&lt;210&gt; 18

&lt;211&gt; 320

&lt;212&gt; PRT

&lt;213&gt; Haematococcus pluvialis

&lt;400&gt; 18

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala  
 1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His  
 20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala  
 35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr  
 50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile  
 65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu  
 85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala  
 100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr  
 115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu  
 130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp  
 145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly  
 165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val  
 180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe  
 195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro  
 210 215 220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala  
 225 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro  
 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr  
 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp  
 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu  
 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala  
 305 310 315 320

<210> 19

<211> 729

<212> DNA

<213> Agrobacterium aurantiacum

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 19  
 atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48  
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
 1 5 10 15  
 atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96  
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
 20 25 30  
 gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca 144  
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
 35 40 45  
 aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192  
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
 50 55 60  
 cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat 240  
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
 65 70 75 80



gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288  
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
 85 90 95

cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc 336  
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
 100 105 110

gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc 384  
 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
 115 120 125

cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc 432  
 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
 130 135 140

gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac 480  
 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
 145 150 155 160

gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc 528  
 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
 165 170 175

gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg 576  
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
 180 185 190

gac cgc cac aat gcg cgg tcg tcg cgg atc agc gac ccc gtg tcg ctg 624  
 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
 195 200 205

ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac 672  
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
 210 215 220

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720  
 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
 225 230 235 240

acc gca tga 729  
 Thr Ala

&lt;210&gt; 20

&lt;211&gt; 242

&lt;212&gt; PRT

&lt;213&gt; Agrobacterium aurantiacum

&lt;400&gt; 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
 225 230 235 240

Thr Ala

<210> 21

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99)..(827)

<223>

<400> 21

```

ctgcaggccg ggccccggtgg ccaatggtcg caaccggcag gactggaaca ggacggcggg      60

ccggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct      116
                               Met Ser Gly Arg Lys Pro
                               1           5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc      164
Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile
                10                15                20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat      212
Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp
                25                30                35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc      260
Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr
                40                45                50

tgg ctg tgc gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg      308
Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly
                55                60                65                70

tcc gtg gtg ccg ggg ccg cgc gcc aat gcg gcg atc ggg caa ctg      356
Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu
                75                80                85

```

gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag 404  
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys  
90 95 100

cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc 452  
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe  
105 110 115

ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat 500  
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr  
120 125 130

ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat 548  
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr  
135 140 145 150

gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc 596  
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val  
155 160 165

ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg 644  
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu  
170 175 180

ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg 692  
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg  
185 190 195

tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc 740  
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe  
200 205 210

ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg 788  
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp  
215 220 225 230

cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct 837  
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala  
235 240

cattgtcgtg gcgacagtcc tcgtgatgga gctgaccgcc tattccgtcc accgctggat 897

tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc 957

gttggaagaag aacgacctct acggcgctcgt cttcgcggtg ctggcgacga tcctcttcac 1017

cgtgggcgcc tattggtggc cgggtgctgtg gtggatcgcc ctgggcatga cgggtctatgg 1077

gttgatctat ttcacctgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat 1137

tcgcgggcgg ggctatttcc gcaggctcta ccaagctcat cgcctgcacc acgcggtcga 1197

ggggcgggac cactgcgtca gcttcggctt catctatgcc ccaccgtgg acaagctgaa 1257

gcaggatctg aagcggtcgg gtgtcctgcg cccccaggac gagcgtccgt cgtgatctct 1317

gatccccggcg tggccgcatg aaatccgacg tgctgctggc aggggcccgc cttgccaacg 1377  
 gactgatcgc gctggcgatc cgcaaggcgc ggcccgcact tcgctgctg ctgctggacc 1437  
 gtgcggcggg cgcctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc 1497  
 actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt 1557  
 tcccagacca ttgcggaagg ctccggggcg gatatggctc gatcgacggg cgggggctga 1617  
 tgcgtgcggt gacc 1631

<210> 22

<211> 242

<212> PRT

<213> *Alcaligenes* sp.

<400> 22

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu  
 1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe  
 20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu  
 35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
 50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
 65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
 85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr  
 100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly  
 115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
 130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
 145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe  
 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro  
 180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu  
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly  
 225 230 235 240

Arg Ala

<210> 23

<211> 729

<212> DNA

<213> *Paracoccus marcusii*

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 23

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu	
1 5 10 15	
atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat	96
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
20 25 30	
gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg	144
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala	
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg	192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	
180 185 190	
gac cgc cat aat gcg cgg tcg tcg ccg atc agc gac cct gtg tcg ctg	624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu	
195 200 205	
ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac	672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His	
210 215 220	
ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac	720

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
 225 230 235 240

acc gca tga  
 Thr Ala

729

&lt;210&gt; 24

&lt;211&gt; 242

&lt;212&gt; PRT

<213> *Paracoccus marcusii*

&lt;400&gt; 24

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
 1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
 20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala  
 35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
 50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
 65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
 85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
 100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
 130 135 140



36

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
 145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
 225 230 235 240

Thr Ala

<210> 25

<211> 1629

<212> DNA

<213> Synechococystis

<220>

<221> CDS

<222> (1)..(1629)

<223>

<400> 25

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48  
 Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu  
 1 5 10 15

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta 96  
 Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu  
 20 25 30

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg	144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met	
35 40 45	
ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac	192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His	
50 55 60	
gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag	240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln	
65 70 75 80	
tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg	288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly	
85 90 95	
ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt	336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys	
100 105 110	
gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa	384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln	
115 120 125	
ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt	432
Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe	
130 135 140	
aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg	480
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp	
145 150 155 160	
gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg	528
Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala	
165 170 175	
ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat	576
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn	
180 185 190	
gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt	624
Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys	
195 200 205	
tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg	672
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met	
210 215 220	
atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga	720
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly	
225 230 235 240	
ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa	768
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln	
245 250 255	

ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa	816
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu	
260 265 270	
aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg	864
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
275 280 285	
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	
290 295 300	
caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg	960
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly	
305 310 315 320	
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa	1008
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys	
325 330 335	
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg	1056
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly	
340 345 350	
ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat	1104
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His	
355 360 365	
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct	1152
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala	
370 375 380	
aat ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg	1200
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met	
385 390 395 400	
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac	1248
Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr	
405 410 415	
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc	1296
Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr	
420 425 430	
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg	1344
Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr	
435 440 445	
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa	1392
Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu	
450 455 460	
agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc	1440
Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val	
465 470 475 480	

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488  
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu  
 485 490 495

cgg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536  
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr  
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584  
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg  
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629  
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp  
 530 535 540

<210> 26

<211> 542

<212> PRT

<213> Synechococystis

<400> 26

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu  
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu  
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met  
 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His  
 50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln  
 65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly  
 85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys  
 100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln  
 115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe  
 130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp  
 145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala  
 165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn  
 180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys  
 195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met  
 210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly  
 225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln  
 245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu  
 260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg  
 275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu  
 290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly  
 305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys  
 325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly  
 340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His  
 355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala  
 370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met  
 385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr  
 405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr  
 420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr  
 435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu  
 450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val  
 465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu  
 485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr  
 500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg  
 515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp  
 530 535 540

&lt;211&gt; 776

&lt;212&gt; DNA

&lt;213&gt; Bradyrhizobium sp.

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (774)

&lt;223&gt;

&lt;400&gt; 27

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc	48
Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg	
1 5 10 15	
gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc	96
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile	
20 25 30	
atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg	144
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro	
35 40 45	
ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag	192
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln	
50 55 60	
acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac	240
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His	
65 70 75 80	
ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag	288
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln	
85 90 95	
ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc	336
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val	
100 105 110	
gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat	384
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp	
115 120 125	
ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt	432
Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe	
130 135 140	
ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc	480

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val  
 145 150 155 160  
 tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg 528  
 Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu  
 165 170 175  
 ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc 576  
 Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr  
 180 185 190  
 ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat 624  
 Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp  
 195 200 205  
 cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg 672  
 Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu  
 210 215 220  
 acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat 720  
 Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp  
 225 230 235 240  
 gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg 768  
 Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg  
 245 250 255  
 cgt gac ta 776  
 Arg Asp

&lt;210&gt; 28

&lt;211&gt; 258

&lt;212&gt; PRT

&lt;213&gt; Bradyrhizobium sp.

&lt;400&gt; 28

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg  
 1 5 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile  
 20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro  
 35 40 45



Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln  
 50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His  
 65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln  
 85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val  
 100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp  
 115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe  
 130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val  
 145 150 155 160

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu  
 165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr  
 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp  
 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu  
 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp  
 225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg  
 245 250 255

Arg Asp

&lt;210&gt; 29

&lt;211&gt; 777

&lt;212&gt; DNA

&lt;213&gt; Nostoc sp.

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(777)

&lt;223&gt;

&lt;400&gt; 29

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta	48
Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu	
1 5 10 15	

ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt	96
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe	
20 25 30	

att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	

ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	

atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	

gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	
85 90 95	

ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa	336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys	
100 105 110	

gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat	384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp	
115 120 125	

tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg	432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp	
130 135 140	

tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga 480  
 Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly  
 145 150 155 160  
  
 tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa 528  
 Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu  
 165 170 175  
  
 aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta 576  
 Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val  
 180 185 190  
  
 caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624  
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly  
 195 200 205  
  
 ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672  
 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe  
 210 215 220  
  
 tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720  
 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His  
 225 230 235 240  
  
 gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768  
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile  
 245 250 255  
  
 tct tta taa 777  
 Ser Leu

&lt;210&gt; 30

&lt;211&gt; 258

&lt;212&gt; PRT

&lt;213&gt; Nostoc sp.

&lt;400&gt; 30

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu  
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe  
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu  
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala  
50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His  
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn  
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys  
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp  
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp  
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly  
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu  
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val  
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly  
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe  
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His  
225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile  
245 250 255

Ser Leu

<210> 31

<211> 831

<212> DNA

<213> Haematococcus pluvialis

**<220>**

<221> CDS

<222> (1) .. (831)

**<223>**

<400> 31

atg	cca	tcc	gag	tcg	tca	gac	gca	gct	cgt	cct	gtg	ttg	aag	cac	gcc	48
Met	Pro	Ser	Glu	Ser	Ser	Asp	Ala	Ala	Arg	Pro	Val	Leu	Lys	His	Ala	
1				5					10					15		

tat	aaa	cct	cca	gca	tct	gac	gcc	aag	ggc	atc	act	atg	gcg	ctg	acc	.96
Tyr	Lys	Pro	Pro	Ala	Ser	Asp	Ala	Lys	Gly	Ile	Thr	Met	Ala	Leu	Thr	
		20						25					30			

atc att ggc acc tgg acc gca gtg ttt tta cac gca ata ttc caa atc 144  
Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile  
35 40 45

agg cta ccg aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa 192  
Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu  
50 55 60

gcc aca gcc cag ctg ttg ggc gga agc agc agc cta ttg cac atc gcc 240  
Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala  
65 70 75 80

gca gtc ttc att gta ctt gag ttt ctg tac act ggt cta ttc atc acc 288  
Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr  
85 90 95

acg cat gat gca atg cat ggc acc ata gct ttg agg aac agg cag ctc 336  
Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg Asn Arg Gln Leu  
100 105 110

aat gat ctc ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac 384  
Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp  
115 120 125

tac agc atg cac tgg gag cac cac aac cat act ggc gaa gtg ggg aaa 432

Tyr	Ser	Met	His	Trp	Glu	His	His	Asn	His	Thr	Gly	Glu	Val	Gly	Lys		
130						135					140						
gac cct gac ttc cac aaa gga aat cct ggc ctt gtc ccc tgg ttc gcc 480																	
Asp	Pro	Asp	Phe	His	Lys	Gly	Asn	Pro	Gly	Leu	Val	Pro	Trp	Phe	Ala		
145					150				155					160			
agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg gca 528																	
Ser	Phe	Met	Ser	Ser	Tyr	Met	Ser	Leu	Trp	Gln	Phe	Ala	Arg	Leu	Ala		
			165					170					175				
tgg tgg gca gtg gtg atg caa acg ttg ggg gcc ccc atg gcg aat ctc 576																	
Trp	Trp	Ala	Val	Val	Met	Gln	Thr	Leu	Gly	Ala	Pro	Met	Ala	Asn	Leu		
			180					185					190				
cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc ttc 624																	
Leu	Val	Phe	Met	Ala	Ala	Ala	Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	Phe		
		195					200					205					
tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca gca 672																	
Tyr	Phe	Gly	Thr	Tyr	Leu	Pro	His	Lys	Pro	Glu	Pro	Gly	Pro	Ala	Ala		
	210					215					220						
ggc tct cag gtc atg tct tgg ttc agg gcc aag aca agt gag gca tct 720																	
Gly	Ser	Gln	Val	Met	Ser	Trp	Phe	Arg	Ala	Lys	Thr	Ser	Glu	Ala	Ser		
225					230					235				240			
gat gtg atg agc ttc ctg aca tgc tac cac ttt gac ctg ttt gcc ccc 768																	
Asp	Val	Met	Ser	Phe	Leu	Thr	Cys	Tyr	His	Phe	Asp	Leu	Phe	Ala	Pro		
				245				250						255			
tgg tgg cag ctg ccc cac tgc cgc cgc ctg tct ggg cgt ggc ctg gtg 816																	
Trp	Trp	Gln	Leu	Pro	His	Cys	Arg	Arg	Leu	Ser	Gly	Arg	Gly	Leu	Val		
			260					265					270				
cct gcc ttg gca tga 831																	
Pro	Ala	Leu	Ala														
			275														

&lt;210&gt; 32

&lt;211&gt; 276

&lt;212&gt; PRT

&lt;213&gt; Haematococcus pluvialis

&lt;400&gt; 32

Met	Pro	Ser	Glu	Ser	Ser	Asp	Ala	Ala	Arg	Pro	Val	Leu	Lys	His	Ala
1				5					10					15	

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr  
 20 25 30

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile  
 35 40 45

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu  
 50 55 60

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala  
 65 70 75 80

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr  
 85 90 95

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg Asn Arg Gln Leu  
 100 105 110

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp  
 115 120 125

Tyr Ser Met His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys  
 130 135 140

Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe Ala  
 145 150 155 160

Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu Ala  
 165 170 175

Trp Trp Ala Val Val Met Gln Thr Leu Gly Ala Pro Met Ala Asn Leu  
 180 185 190

Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe  
 195 200 205

Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala Ala  
 210 215 220

Gly Ser Gln Val Met Ser Trp Phe Arg Ala Lys Thr Ser Glu Ala Ser  
 225 230 235 240

Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu Phe Ala Pro  
 245 250 255

Trp Trp Gln Leu Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val  
 260 265 270

Pro Ala Leu Ala  
 275

<210> 33

<211> 729

<212> DNA

<213> Paracoccus sp. MBIC1143

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 33

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48  
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96  
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
 20 25 30

gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca 144  
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192  
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
 50 55 60

cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat 240  
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
 65 70 75 80

gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288  
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
 85 90 95



cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc 336  
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
 100 105 110

gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc 384  
 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
 115 120 125

cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc 432  
 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
 130 135 140

gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac 480  
 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
 145 150 155 160

gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc 528  
 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
 165 170 175

gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg 576  
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
 180 185 190

gac cgc cac aat gcg ccg tcg tcg ccg atc agc gac ccc gtg tcg ctg 624  
 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
 195 200 205

ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac 672  
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
 210 215 220

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720  
 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
 225 230 235 240

acc gca tga 729  
 Thr Ala

&lt;210&gt; 34

&lt;211&gt; 242

&lt;212&gt; PRT

&lt;213&gt; Paracoccus sp. MBIC1143

&lt;400&gt; 34

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
 1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
225 230 235 240

Thr Ala

&lt;210&gt; 35

&lt;211&gt; 735

&lt;212&gt; DNA

&lt;213&gt; Brevundimonas aurantiaca

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (735)

&lt;223&gt;

&lt;400&gt; 35

atg acc gcc gcc gtc gcc gag cca cgc acc gtc ccg cgc cag acc tgg	48
Met Thr Ala Ala Val Ala Glu Pro Arg Thr Val Pro Arg Gln Thr Trp	
1 5 10 15	

atc ggt ctg acc ctg gcg gga atg atc gtg gcg gga tgg gcg gtt ctg	96
Ile Gly Leu Thr Leu Ala Gly Met Ile Val Ala Gly Trp Ala Val Leu	
20 25 30	

cat gtc tac ggc gtc tat ttt cac cga tgg ggg ccg ttg acc ctg gtg	144
His Val Tyr Gly Val Tyr Phe His Arg Trp Gly Pro Leu Thr Leu Val	
35 40 45	

atc gcc ccg gcg atc gtg gcg gtc cag acc tgg ttg tcg gtc ggc ctt	192
Ile Ala Pro Ala Ile Val Ala Val Gln Thr Trp Leu Ser Val Gly Leu	
50 55 60	

ttc atc gtc gcc cat gac gcc atg tac ggc tcc ctg gcg ccg gga cgg	240
Phe Ile Val Ala His Asp Ala Met Tyr Gly Ser Leu Ala Pro Gly Arg	
65 70 75 80	

ccg ccg ctg aac gcc gca gtc ggc ccg ctg acc ctg ggg ctc tat gcg	288
Pro Arg Leu Asn Ala Ala Val Gly Arg Leu Thr Leu Gly Leu Tyr Ala	
85 90 95	

ggc ttc cgc ttc gat ccg ctg aag acg gcg cac cac gcc cac cac gcc	336
Gly Phe Arg Phe Asp Arg Leu Lys Thr Ala His His Ala His His Ala	
100 105 110	

gcg ccc gcc acg gcc gac gac ccg gat ttt cac gcc ccg gcg ccc cgc	384
---	-----

Ala Pro Gly Thr Ala Asp Asp Pro Asp Phe His Ala Pro Ala Pro Arg	
115 120 125	
gcc ttc ctt ccc tgg ttc ctg aac ttc ttt cgc acc tat ttc ggc tgg	432
Ala Phe Leu Pro Trp Phe Leu Asn Phe Phe Arg Thr Tyr Phe Gly Trp	
130 135 140	
cgc gag atg gcg gtc ctg acc gcc ctg gtc ctg atc gcc ctc ttc ggc	480
Arg Glu Met Ala Val Leu Thr Ala Leu Val Leu Ile Ala Leu Phe Gly	
145 150 155 160	
ctg ggg gcg cgg ccg gcc aat ctc ctg acc ttc tgg gcc gcg ccg gcc	528
Leu Gly Ala Arg Pro Ala Asn Leu Leu Thr Phe Trp Ala Ala Pro Ala	
165 170 175	
ctg ctt tca gcg ctt cag ctc ttc acc ttc ggc acc tgg ctg ccg cac	576
Leu Leu Ser Ala Leu Gln Leu Phe Thr Phe Gly Thr Trp Leu Pro His	
180 185 190	
cgc cac acc gac cag ccg ttc gcc gac gcg cac cac gcc cgc agc agc	624
Arg His Thr Asp Gln Pro Phe Ala Asp Ala His His Ala Arg Ser Ser	
195 200 205	
ggc tac ggc ccc gtg ctt tcc ctg ctc acc tgt ttc cac ttc ggc cgc	672
Gly Tyr Gly Pro Val Leu Ser Leu Leu Thr Cys Phe His Phe Gly Arg	
210 215 220	
cac cac gaa cac cat ctg agc ccc tgg cgg ccc tgg tgg cgt ctg tgg	720
His His Glu His His Leu Ser Pro Trp Arg Pro Trp Trp Arg Leu Trp	
225 230 235 240	
cgc ggc gag tct tga	735
Arg Gly Glu Ser	

&lt;210&gt; 36

&lt;211&gt; 244

&lt;212&gt; PRT

&lt;213&gt; Brevundimonas aurantiaca

&lt;400&gt; 36

Met Thr Ala Ala Val Ala Glu Pro Arg Thr Val Pro Arg Gln Thr Trp
1 5 10 15
Ile Gly Leu Thr Leu Ala Gly Met Ile Val Ala Gly Trp Ala Val Leu
20 25 30

His Val Tyr Gly Val Tyr Phe His Arg Trp Gly Pro Leu Thr Leu Val  
 35 40 45

Ile Ala Pro Ala Ile Val Ala Val Gln Thr Trp Leu Ser Val Gly Leu  
 50 55 60

Phe Ile Val Ala His Asp Ala Met Tyr Gly Ser Leu Ala Pro Gly Arg  
 65 70 75 80

Pro Arg Leu Asn Ala Ala Val Gly Arg Leu Thr Leu Gly Leu Tyr Ala  
 85 90 95

Gly Phe Arg Phe Asp Arg Leu Lys Thr Ala His His Ala His His Ala  
 100 105 110

Ala Pro Gly Thr Ala Asp Asp Pro Asp Phe His Ala Pro Ala Pro Arg  
 115 120 125

Ala Phe Leu Pro Trp Phe Leu Asn Phe Phe Arg Thr Tyr Phe Gly Trp  
 130 135 140

Arg Glu Met Ala Val Leu Thr Ala Leu Val Leu Ile Ala Leu Phe Gly  
 145 150 155 160

Leu Gly Ala Arg Pro Ala Asn Leu Leu Thr Phe Trp Ala Ala Pro Ala  
 165 170 175

Leu Leu Ser Ala Leu Gln Leu Phe Thr Phe Gly Thr Trp Leu Pro His  
 180 185 190

Arg His Thr Asp Gln Pro Phe Ala Asp Ala His His Ala Arg Ser Ser  
 195 200 205

Gly Tyr Gly Pro Val Leu Ser Leu Leu Thr Cys Phe His Phe Gly Arg  
 210 215 220

His His Glu His His Leu Ser Pro Trp Arg Pro Trp Trp Arg Leu Trp  
 225 230 235 240

Arg Gly Glu Ser

&lt;210&gt; 37

&lt;211&gt; 690

&lt;212&gt; DNA

&lt;213&gt; Nodularia spumigena NSOR10

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(690)

&lt;223&gt;

&lt;400&gt; 37

atg	gcg	atc	gcc	att	att	agt	ata	tgg	gct	atc	agc	cta	ggg	ttg	tta	48
Met	Ala	Ile	Ala	Ile	Ile	Ser	Ile	Trp	Ala	Ile	Ser	Leu	Gly	Leu	Leu	
1			5					10					15			

ctt	tat	att	gat	ata	tcc	caa	ttc	aag	ttt	tgg	atg	ttg	tta	ccg	ctc	96
Leu	Tyr	Ile	Asp	Ile	Ser	Gln	Phe	Lys	Phe	Trp	Met	Leu	Leu	Pro	Leu	
			20					25					30			

ata	ttt	tgg	caa	aca	ttt	tta	tat	acg	gga	tta	ttt	att	aca	gct	cat	144
Ile	Phe	Trp	Gln	Thr	Phe	Leu	Tyr	Thr	Gly	Leu	Phe	Ile	Thr	Ala	His	
			35					40					45			

gat	gcc	atg	cat	ggg	gta	gtt	ttt	ccc	aaa	aat	ccc	aaa	atc	aac	cat	192
Asp	Ala	Met	His	Gly	Val	Val	Phe	Pro	Lys	Asn	Pro	Lys	Ile	Asn	His	
			50				55					60				

ttc	att	ggc	tca	ttg	tgc	ctg	ttt	ctt	tat	ggg	ctt	tta	cct	tat	caa	240
Phe	Ile	Gly	Ser	Leu	Cys	Leu	Phe	Leu	Tyr	Gly	Leu	Leu	Pro	Tyr	Gln	
65					70					75					80	

aaa	ctt	tta	aaa	aag	cat	tgg	cta	cat	cac	cat	aat	cca	gcc	agt	gaa	288
Lys	Leu	Leu	Lys	Lys	His	Trp	Leu	His	His	His	Asn	Pro	Ala	Ser	Glu	
				85					90						95	

aca	gat	cca	gat	ttt	cac	aac	ggg	aag	cag	aaa	aac	ttt	ttt	gct	tgg	336
Thr	Asp	Pro	Asp	Phe	His	Asn	Gly	Lys	Gln	Lys	Asn	Phe	Phe	Ala	Trp	
				100				105						110		

tat	tta	tat	ttt	atg	aag	cgt	tac	tgg	agt	tgg	tta	caa	att	atc	aca	384
Tyr	Leu	Tyr	Phe	Met	Lys	Arg	Tyr	Trp	Ser	Trp	Leu	Gln	Ile	Ile	Thr	
			115				120					125				

tta	atg	att	att	tat	aac	tta	cta	aaa	tat	ata	tgg	cat	ttt	cca	gag	432
Leu	Met	Ile	Ile	Tyr	Asn	Leu	Leu	Lys	Tyr	Ile	Trp	His	Phe	Pro	Glu	
				130				135				140				

gat aat atg act tat ttt tgg gta gtt ccc tca att tta agt tct tta 480  
 Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu  
 145 150 155 160

caa tta ttt tat ttt gga act ttt cta ccc cac agt gag cct gta gaa 528  
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu  
 165 170 175

ggg tat aaa gag cct cat cgt tcc caa act att agc cgt ccc att tgg 576  
 Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp  
 180 185 190

tgg tca ttt ata act tgt tac cat ttt ggt tat cat tac gaa cat cat 624  
 Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His  
 195 200 205

gaa tac ccc cat gtt cct tgg tgg caa tta cca gaa att tat aaa atg 672  
 Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met  
 210 215 220

tct aaa tca aat ttg tga 690  
 Ser Lys Ser Asn Leu  
 225

<210> 38

<211> 229

<212> PRT

<213> Nodularia spumigena NSOR10

<400> 38

Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu  
 1 5 10 15

Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu  
 20 25 30

Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His  
 35 40 45

Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His  
 50 55 60

Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu Leu Pro Tyr Gln  
 65 70 75 80

Lys Leu Leu Lys Lys His Trp Leu His His His Asn Pro Ala Ser Glu  
85 90 95

Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn Phe Phe Ala Trp  
100 105 110

Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu Gln Ile Ile Thr  
115 120 125

Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp His Phe Pro Glu  
130 135 140

Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu  
145 150 155 160

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu  
165 170 175

Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp  
180 185 190

Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His  
195 200 205

Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met  
210 215 220

Ser Lys Ser Asn Leu  
225

<210> 39

<211> 789

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> CDS



&lt;222&gt; (1)..(789)

&lt;223&gt;

&lt;400&gt; 39

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa	48
Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln	
1 5 10 15	
tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta	96
Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val	
20 25 30	
att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat	144
Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn	
35 40 45	
tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa	192
Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln	
50 55 60	
atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat	240
Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His	
65 70 75 80	
ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca	288
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser	
85 90 95	
cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag	336
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys	
100 105 110	
aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat	384
Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp	
115 120 125	
ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc	432
Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe	
130 135 140	
atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta	480
Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu	
145 150 155 160	
ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc	528
Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile	
165 170 175	
tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat	576
Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr	
180 185 190	
ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat	624

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
 195 200 205

ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc 672  
 Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
 210 215 220

gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat 720  
 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
 225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768  
 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
 245 250 255

aat tca gta acc aat tcg taa 789  
 Asn Ser Val Thr Asn Ser  
 260

<210> 40

<211> 262

<212> PRT

<213> Nostoc punctiforme ATCC 29133

<400> 40

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
 1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
 20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
 35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
 50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
 65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
 85 90 95

62

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
 100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp  
 115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
 130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
 145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
 165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
 180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
 195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
 210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
 225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
 245 250 255

Asn Ser Val Thr Asn Ser  
 260

<210> 41

<211> 762

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

&lt;221&gt; CDS

&lt;222&gt; (1)..(762)

&lt;223&gt;

&lt;400&gt; 41

gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca	48
Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro	
1 5 10 15	

gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc	96
Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val	
20 25 30	

att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac	144
Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp	
35 40 45	

atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa	192
Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln	
50 55 60	

aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat	240
Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His	
65 70 75 80	

ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca	288
Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr	
85 90 95	

ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa	336
Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys	
100 105 110	

aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat	384
Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp	
115 120 125	

ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt	432
Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe	
130 135 140	

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att	480
Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile	
145 150 155 160	

tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act	528
Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr	
165 170 175	

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat	576
Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr	
180 185 190	

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624  
 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
 195 200 205  
 cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672  
 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
 210 215 220  
 acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720  
 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
 225 230 235 240  
 att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762  
 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
 245 250

&lt;210&gt; 42

&lt;211&gt; 253

&lt;212&gt; PRT

&lt;213&gt; Nostoc punctiforme ATCC 29133

&lt;400&gt; 42

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro  
 1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val  
 20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp  
 35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln  
 50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His  
 65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr  
 85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys  
 100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp  
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe  
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile  
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr  
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 43

<211> 1536

<212> DNA

<213> Deinococcus radiodurans R1

<220>

<221> CDS

<222> (1)..(1536)

<223>

<400> 43  
 atg ccg gat tac gac ctg atc gtc atg ggc gcg ggc cac aac gcg ctg 48  
 Met Pro Asp Tyr Asp Leu Ile Val Met Gly Ala Gly His Asn Ala Leu  
 1 5 10 15

gtg act gct gcc tac gcc gcc cgg gcg ggc ctg aaa gtc ggc gtg ttc 96  
 Val Thr Ala Ala Tyr Ala Ala Arg Ala Gly Leu Lys Val Gly Val Phe  
 20 25 30

gag cgg cgg cac ctc gtc ggc ggg gcg gtc agc acc gag gag gtc gtg 144  
 Glu Arg Arg His Leu Val Gly Gly Ala Val Ser Thr Glu Glu Val Val  
 35 40 45

ccc ggt tac cgc ttc gac tac ggc ggc agc gcc cac atc ctg att cgg 192  
 Pro Gly Tyr Arg Phe Asp Tyr Gly Gly Ser Ala His Ile Leu Ile Arg  
 50 55 60

atg acg ccc atc gtg cgc gaa ctc gaa ctc acg cgg cac ggg ctg cat 240  
 Met Thr Pro Ile Val Arg Glu Leu Glu Leu Thr Arg His Gly Leu His  
 65 70 75 80

tac ctc gaa gtg gac cct atg ttt cac gct tcc gac ggt gaa acg ccc 288  
 Tyr Leu Glu Val Asp Pro Met Phe His Ala Ser Asp Gly Glu Thr Pro  
 85 90 95

tgg ttc att cac cgc gac gcc ggg cgg acc atc cgc gaa ctg gac gaa 336  
 Trp Phe Ile His Arg Asp Ala Gly Arg Thr Ile Arg Glu Leu Asp Glu  
 100 105 110

aag ttt ccc ggg cag ggc gac gcc tac ggg cgc ttt ctc gac gat tgg 384  
 Lys Phe Pro Gly Gln Gly Asp Ala Tyr Gly Arg Phe Leu Asp Asp Trp  
 115 120 125

aca ccc ttc gcg cgc gcc gtg gcc gac ctg ttc aac tcg gcg ccg ggg 432  
 Thr Pro Phe Ala Arg Ala Val Ala Asp Leu Phe Asn Ser Ala Pro Gly  
 130 135 140

ccg ctc gac ctg ggc aaa atg gtg atg cgc agc ggc cag ggc aag gac 480  
 Pro Leu Asp Leu Gly Lys Met Val Met Arg Ser Gly Gln Gly Lys Asp  
 145 150 155 160

tgg aac gag cag ctc ccg cgc atc ctg cgg ccc tac ggc gac gtg gcg 528  
 Trp Asn Glu Gln Leu Pro Arg Ile Leu Arg Pro Tyr Gly Asp Val Ala  
 165 170 175

cgc gag tac ttc agc gag gag cgc gtg cgg gct ccc ctg acc tgg atg 576  
 Arg Glu Tyr Phe Ser Glu Glu Arg Val Arg Ala Pro Leu Thr Trp Met  
 180 185 190

gcg gcc cag agc ggc ccc cca ccc tcg gac ccg ctg agc gcg ccc ttt 624  
 Ala Ala Gln Ser Gly Pro Pro Pro Ser Asp Pro Leu Ser Ala Pro Phe  
 195 200 205

ttg ctg tgg cac ccg ctc tac cac gaa ggc ggc gtg gcg cgg ccc aaa 672

Leu Leu Trp His Pro Leu Tyr His Glu Gly Gly Val Ala Arg Pro Lys	
210 215 220	
ggc ggc agc ggc ggc ctg acc aaa gcc ctg cgc cgg gcc acc gag gcc	720
Gly Gly Ser Gly Gly Leu Thr Lys Ala Leu Arg Arg Ala Thr Glu Ala	
225 230 235 240	
gaa ggc ggc gag gtc ttc acc gac gcg ccg gtc aag gaa att ctg gtc	768
Glu Gly Gly Glu Val Phe Thr Asp Ala Pro Val Lys Glu Ile Leu Val	
245 250 255	
aag gac ggc aag gcg cag ggc atc cgg ctg gaa agc ggc gag acg tac	816
Lys Asp Gly Lys Ala Gln Gly Ile Arg Leu Glu Ser Gly Glu Thr Tyr	
260 265 270	
acc gcc cgc gcc gtc gtg tcg ggc gtc cac atc ctg acc act gcg aat	864
Thr Ala Arg Ala Val Val Ser Gly Val His Ile Leu Thr Thr Ala Asn	
275 280 285	
gcc ctg ccc gcc gaa tat gtc cct agc gcc gcc agg aat gtg cgc gtg	912
Ala Leu Pro Ala Glu Tyr Val Pro Ser Ala Ala Arg Asn Val Arg Val	
290 295 300	
ggc aac ggc ttc ggc atg att ttg cgc ctc gcc ctc agt gaa aaa gtc	960
Gly Asn Gly Phe Gly Met Ile Leu Arg Leu Ala Leu Ser Glu Lys Val	
305 310 315 320	
aaa tac cgt cac cac acc gag ccc gac tca cgc atc ggc ctg gga ttg	1008
Lys Tyr Arg His His Thr Glu Pro Asp Ser Arg Ile Gly Leu Gly Leu	
325 330 335	
ctg atc aaa aac gag cgg caa atc atg cag ggc tac ggc gaa tac ctc	1056
Leu Ile Lys Asn Glu Arg Gln Ile Met Gln Gly Tyr Gly Glu Tyr Leu	
340 345 350	
gcc ggg cag ccc acc acc gac ccg ccc ctc gtc gcc atg agc ttc agc	1104
Ala Gly Gln Pro Thr Thr Asp Pro Pro Leu Val Ala Met Ser Phe Ser	
355 360 365	
gcg gtg gac gac tcg ctc gcc cca ccg aac ggc gac gtg ttg tgg ctg	1152
Ala Val Asp Asp Ser Leu Ala Pro Pro Asn Gly Asp Val Leu Trp Leu	
370 375 380	
tgg gcg cag tac tac ccc ttc gag ctc gcc acc ggg agc tgg gaa acg	1200
Trp Ala Gln Tyr Tyr Pro Phe Glu Leu Ala Thr Gly Ser Trp Glu Thr	
385 390 395 400	
cgc acc gcc gaa gcg cgg gag aac atc ctg cgg gcc ttt gag cac tac	1248
Arg Thr Ala Glu Ala Arg Glu Asn Ile Leu Arg Ala Phe Glu His Tyr	
405 410 415	
gcg ccg ggc acc cgc gac acg att gtg ggc gaa ctc gtg cag acg ccg	1296
Ala Pro Gly Thr Arg Asp Thr Ile Val Gly Glu Leu Val Gln Thr Pro	
420 425 430	
cag tgg ctg gaa acc aac ctc ggc ctg cac cgg ggc aac gtg atg cac	1344



Gln Trp Leu Glu Thr Asn Leu Gly Leu His Arg Gly Asn Val Met His  
 435 440 445

ctg gaa atg tcc ttc gac cag atg ttc tcc ttc cgc ccc tgg ctg aaa 1392  
 Leu Glu Met Ser Phe Asp Gln Met Phe Ser Phe Arg Pro Trp Leu Lys  
 450 455 460

gcg agc cag tac cgc tgg ccg ggc gtg cag ggg ctg tac ctc acc ggc 1440  
 Ala Ser Gln Tyr Arg Trp Pro Gly Val Gln Gly Leu Tyr Leu Thr Gly  
 465 470 475 480

gcc agc acc cac ccc ggc gga ggc atc atg ggc gcc tcg gga cgc aac 1488  
 Ala Ser Thr His Pro Gly Gly Gly Ile Met Gly Ala Ser Gly Arg Asn  
 485 490 495

gcg gcg cgg gtc atc gtg aag gac ctg acg cgg agg cgc tgg aaa tga 1536  
 Ala Ala Arg Val Ile Val Lys Asp Leu Thr Arg Arg Arg Trp Lys  
 500 505 510

&lt;210&gt; 44

&lt;211&gt; 511

&lt;212&gt; PRT

&lt;213&gt; Deinococcus radiodurans R1

&lt;400&gt; 44

Met Pro Asp Tyr Asp Leu Ile Val Met Gly Ala Gly His Asn Ala Leu  
 1 5 10 15

Val Thr Ala Ala Tyr Ala Ala Arg Ala Gly Leu Lys Val Gly Val Phe  
 20 25 30

Glu Arg Arg His Leu Val Gly Gly Ala Val Ser Thr Glu Glu Val Val  
 35 40 45

Pro Gly Tyr Arg Phe Asp Tyr Gly Gly Ser Ala His Ile Leu Ile Arg  
 50 55 60

Met Thr Pro Ile Val Arg Glu Leu Glu Leu Thr Arg His Gly Leu His  
 65 70 75 80

Tyr Leu Glu Val Asp Pro Met Phe His Ala Ser Asp Gly Glu Thr Pro  
 85 90 95

Trp Phe Ile His Arg Asp Ala Gly Arg Thr Ile Arg Glu Leu Asp Glu  
 100 105 110

Lys Phe Pro Gly Gln Gly Asp Ala Tyr Gly Arg Phe Leu Asp Asp Trp  
 115 120 125

Thr Pro Phe Ala Arg Ala Val Ala Asp Leu Phe Asn Ser Ala Pro Gly  
 130 135 140

Pro Leu Asp Leu Gly Lys Met Val Met Arg Ser Gly Gln Gly Lys Asp  
 145 150 155 160

Trp Asn Glu Gln Leu Pro Arg Ile Leu Arg Pro Tyr Gly Asp Val Ala  
 165 170 175

Arg Glu Tyr Phe Ser Glu Glu Arg Val Arg Ala Pro Leu Thr Trp Met  
 180 185 190

Ala Ala Gln Ser Gly Pro Pro Pro Ser Asp Pro Leu Ser Ala Pro Phe  
 195 200 205

Leu Leu Trp His Pro Leu Tyr His Glu Gly Gly Val Ala Arg Pro Lys  
 210 215 220

Gly Gly Ser Gly Gly Leu Thr Lys Ala Leu Arg Arg Ala Thr Glu Ala  
 225 230 235 240

Glu Gly Gly Glu Val Phe Thr Asp Ala Pro Val Lys Glu Ile Leu Val  
 245 250 255

Lys Asp Gly Lys Ala Gln Gly Ile Arg Leu Glu Ser Gly Glu Thr Tyr  
 260 265 270

Thr Ala Arg Ala Val Val Ser Gly Val His Ile Leu Thr Thr Ala Asn  
 275 280 285

Ala Leu Pro Ala Glu Tyr Val Pro Ser Ala Ala Arg Asn Val Arg Val  
 290 295 300

Gly Asn Gly Phe Gly Met Ile Leu Arg Leu Ala Leu Ser Glu Lys Val  
 305 310 315 320

Lys Tyr Arg His His Thr Glu Pro Asp Ser Arg Ile Gly Leu Gly Leu  
 325 330 335

Leu Ile Lys Asn Glu Arg Gln Ile Met Gln Gly Tyr Gly Glu Tyr Leu  
 340 345 350

Ala Gly Gln Pro Thr Thr Asp Pro Pro Leu Val Ala Met Ser Phe Ser  
 355 360 365

Ala Val Asp Asp Ser Leu Ala Pro Pro Asn Gly Asp Val Leu Trp Leu  
 370 375 380

Trp Ala Gln Tyr Tyr Pro Phe Glu Leu Ala Thr Gly Ser Trp Glu Thr  
 385 390 395 400

Arg Thr Ala Glu Ala Arg Glu Asn Ile Leu Arg Ala Phe Glu His Tyr  
 405 410 415

Ala Pro Gly Thr Arg Asp Thr Ile Val Gly Glu Leu Val Gln Thr Pro  
 420 425 430

Gln Trp Leu Glu Thr Asn Leu Gly Leu His Arg Gly Asn Val Met His  
 435 440 445

Leu Glu Met Ser Phe Asp Gln Met Phe Ser Phe Arg Pro Trp Leu Lys  
 450 455 460

Ala Ser Gln Tyr Arg Trp Pro Gly Val Gln Gly Leu Tyr Leu Thr Gly  
 465 470 475 480

Ala Ser Thr His Pro Gly Gly Gly Ile Met Gly Ala Ser Gly Arg Asn  
 485 490 495

Ala Ala Arg Val Ile Val Lys Asp Leu Thr Arg Arg Arg Trp Lys  
 500 505 510

<210> 45

<211> 1650

<212> DNA

<213> Lycopersicon esculentum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (112)..(1614)

&lt;223&gt;

&lt;400&gt; 45

```

ggcaccgagga aacttttctc tcttcactag ctgtttacat gcttgaaatt tcaagatttt      60

aggaccccat ttgaagtttt cttgaaacaa atattaccct gttggaaaaa g atg gat      117
                                   Met Asp
                                   1

act ttg ttg aaa acc cca aat aac ctt gaa ttt ctg aac cca cat cat      165
Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro His His
      5                               10                               15

ggt ttt gct gtt aaa gct agt acc ttt aga tct gag aag cat cat aat      213
Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His His Asn
      20                               25                               30

ttt ggt tct agg aag ttt tgt gaa act ttg ggt aga agt gtt tgt gtt      261
Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val Cys Val
      35                               40                               45                               50

aag ggt agt agt agt gct ctt tta gag ctt gta cct gag acc aaa aag      309
Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys
      55                               60                               65

gag aat ctt gat ttt gag ctt cct atg tat gac cct tca aaa ggg gtt      357
Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val
      70                               75                               80

gtt gtg gat ctt gct gtg gtt ggt ggt ggc cct gca gga ctt gct gtt      405
Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu Ala Val
      85                               90                               95

gca cag caa gtt tct gaa gca gga ctc tct gtt tgt tca att gat ccg      453
Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro
      100                               105                               110

aat cct aaa ttg ata tgg cct aat aac tat ggt gtt tgg gtg gat gaa      501
Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu
      115                               120                               125                               130

ttt gag gct atg gac ttg tta gat tgt cta gat gct acc tgg tct ggt      549
Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly
      135                               140                               145

gca gca gtg tac att gat gat aat acg gct aaa gat ctt cat aga cct      597

```

Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His Arg Pro	
150 155 160	
tat gga agg gtt aac cgg aaa cag ctg aaa tcg aaa atg atg cag aaa	645
Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys	
165 170 175	
tgt ata atg aat ggt gtt aaa ttc cac caa gcc aaa gtt ata aag gtg	693
Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile Lys Val	
180 185 190	
att cat gag gaa tcg aaa tcc atg ttg ata tgc aat gat ggt att act	741
Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly Ile Thr	
195 200 205 210	
att cag gca acg gtg gtg ctc gat gca act ggc ttc tct aga tct ctt	789
Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu	
215 220 225	
gtt cag tat gat aag cct tat aac ccc ggg tat caa gtt gct tat ggc	837
Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly	
230 235 240	
att ttg gct gaa gtg gaa gag cac ccc ttt gat gta aac aag atg gtt	885
Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys Met Val	
245 250 255	
ttc atg gat tgg cga gat tct cat ttg aag aac aat act gat ctc aag	933
Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp Leu Lys	
260 265 270	
gag aga aat agt aga ata cca act ttt ctt tat gca atg cca ttt tca	981
Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser	
275 280 285 290	
tcc aac agg ata ttt ctt gaa gaa aca tca ctc gta gct cgt cct ggc	1029
Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly	
295 300 305	
ttg cgt ata gat gat att caa gaa cga atg gtg gct cgt tta aac cat	1077
Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu Asn His	
310 315 320	
ttg ggg ata aaa gtg aag agc att gaa gaa gat gaa cat tgt cta ata	1125
Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Leu Ile	
325 330 335	
cca atg ggt ggt cca ctt cca gta tta cct cag aga gtc gtt gga atc	1173
Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile	
340 345 350	
ggg ggt aca gct ggc atg gtt cat cca tcc acc ggt tat atg gtg gca	1221
Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala	
355 360 365 370	
agg aca cta gct gcg gct cct gtt gtt gcc aat gcc ata att caa tac	1269

Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr	
375	380 385
ctc ggt tct gaa aga agt cat tcg ggt aat gaa tta tcc aca gct gtt	1317
Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr Ala Val	
390	395 400
tgg aaa gat ttg tgg cct ata gag agg aga cgt caa aga gag ttc ttc	1365
Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe	
405	410 415
tgc ttc ggt atg gat att ctt ctg aag ctt gat tta cct gct aca aga	1413
Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg	
420	425 430
agg ttc ttt gat gca ttc ttt gac tta gaa cct cgt tat tgg cat ggc	1461
Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly	
435	440 445 450
ttc tta tcg tct cga ttg ttt cta cct gaa ctc ata gtt ttt ggg ctg	1509
Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu	
455	460 465
tct cta ttc tct cat gct tca aat act tct aga ttt gag ata atg aca	1557
Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile Met Thr	
470	475 480
aag gga act gtt cca tta gta aat atg atc aac aat ttg tta cag gat	1605
Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu Gln Asp	
485	490 495
aaa gaa tga atccgagtaa ttcggaatct tgtccaatct cgtgcc	1650
Lys Glu	
500	

&lt;210&gt; 46

&lt;211&gt; 500

&lt;212&gt; PRT

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 46

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro
1 5 10 15

His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His
20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val  
 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr  
 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys  
 65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu  
 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile  
 100 105 110

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val  
 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp  
 130 135 140

Ser Gly Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His  
 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met  
 165 170 175

Gln Lys Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile  
 180 185 190

Lys Val Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly  
 195 200 205

Ile Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg  
 210 215 220

Ser Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala  
 225 230 235 240

Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys  
 245 250 255

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp  
 260 265 270

Leu Lys Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro  
 275 280 285

Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg  
 290 295 300

Pro Gly Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu  
 305 310 315 320

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys  
 325 330 335

Leu Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val  
 340 345 350

Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met  
 355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile  
 370 375 380

Gln Tyr Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr  
 385 390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu  
 405 410 415

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala  
 420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp  
 435 440 445

His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe  
 450 455 460

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile  
 465 470 475 480



Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu  
 485 490 495

Gln Asp Lys Glu  
 500

<210> 47

<211> 1666

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(1494)

<223>

<400> 47

atg gaa gct ctt ctc aag cct ttt cca tct ctt tta ctt tcc tct cct 48  
 Met Glu Ala Leu Leu Lys Pro Phe Pro Ser Leu Leu Leu Ser Ser Pro  
 1 5 10 15

aca ccc cat agg tct att ttc caa caa aat ccc tct ttt cta agt ccc 96  
 Thr Pro His Arg Ser Ile Phe Gln Gln Asn Pro Ser Phe Leu Ser Pro  
 20 25 30

acc acc aaa aaa aaa tca aga aaa tgt ctt ctt aga aac aaa agt agt 144  
 Thr Thr Lys Lys Lys Ser Arg Lys Cys Leu Leu Arg Asn Lys Ser Ser  
 35 40 45

aaa ctt ttt tgt agc ttt ctt gat tta gca ccc aca tca aag cca gag 192  
 Lys Leu Phe Cys Ser Phe Leu Asp Leu Ala Pro Thr Ser Lys Pro Glu  
 50 55 60

tct tta gat gtt aac atc tca tgg gtt gat cct aat tcg aat cgg gct 240  
 Ser Leu Asp Val Asn Ile Ser Trp Val Asp Pro Asn Ser Asn Arg Ala  
 65 70 75 80

caa ttc gac gtg atc att atc gga gct ggc cct gct ggg ctc agg cta 288  
 Gln Phe Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Arg Leu  
 85 90 95

gct gaa caa gtt tct aaa tat ggt att aag gta tgt tgt gtt gac cct 336  
 Ala Glu Gln Val Ser Lys Tyr Gly Ile Lys Val Cys Cys Val Asp Pro  
 100 105 110

tca cca ctc tcc atg tgg cca aat aat tat ggt gtt tgg gtt gat gag	384
Ser Pro Leu Ser Met Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu	
115 120 125	
ttt gag aat tta gga ctg gaa aat tgt tta gat cat aaa tgg cct atg	432
Phe Glu Asn Leu Gly Leu Glu Asn Cys Leu Asp His Lys Trp Pro Met	
130 135 140	
act tgt gtg cat ata aat gat aac aaa act aag tat ttg gga aga cca	480
Thr Cys Val His Ile Asn Asp Asn Lys Thr Lys Tyr Leu Gly Arg Pro	
145 150 155 160	
tat ggt aga gtt agt aga aag aag ctg aag ttg aaa ttg ttg aat agt	528
Tyr Gly Arg Val Ser Arg Lys Lys Leu Lys Leu Lys Leu Leu Asn Ser	
165 170 175	
tgt gtt gag aac aga gtg aag ttt tat aaa gct aag gtt tgg aaa gtg	576
Cys Val Glu Asn Arg Val Lys Phe Tyr Lys Ala Lys Val Trp Lys Val	
180 185 190	
gaa cat gaa gaa ttt gag tct tca att gtt tgt gat gat ggt aag aag	624
Glu His Glu Glu Phe Glu Ser Ser Ile Val Cys Asp Asp Gly Lys Lys	
195 200 205	
ata aga ggt agt ttg gtt gtg gat gca agt ggt ttt gct agt gat ttt	672
Ile Arg Gly Ser Leu Val Val Asp Ala Ser Gly Phe Ala Ser Asp Phe	
210 215 220	
ata gag tat gac agg cca aga aac cat ggt tat caa att gct cat ggg	720
Ile Glu Tyr Asp Arg Pro Arg Asn His Gly Tyr Gln Ile Ala His Gly	
225 230 235 240	
gtt tta gta gaa gtt gat aat cat cca ttt gat ttg gat aaa atg gtg	768
Val Leu Val Glu Val Asp Asn His Pro Phe Asp Leu Asp Lys Met Val	
245 250 255	
ctt atg gat tgg agg gat tct cat ttg ggt aat gag cca tat tta agg	816
Leu Met Asp Trp Arg Asp Ser His Leu Gly Asn Glu Pro Tyr Leu Arg	
260 265 270	
gtg aat aat gct aaa gaa cca aca ttc ttg tat gca atg cca ttt gat	864
Val Asn Asn Ala Lys Glu Pro Thr Phe Leu Tyr Ala Met Pro Phe Asp	
275 280 285	
aga gat ttg gtt ttc ttg gaa gag act tct ttg gtg agt cgt cct gtt	912
Arg Asp Leu Val Phe Leu Glu Glu Thr Ser Leu Val Ser Arg Pro Val	
290 295 300	
tta tcg tat atg gaa gta aaa aga agg atg gtg gca aga tta agg cat	960
Leu Ser Tyr Met Glu Val Lys Arg Arg Met Val Ala Arg Leu Arg His	
305 310 315 320	
ttg ggg atc aaa gtg aaa agt gtt att gag gaa gag aaa tgt gtg atc	1008
Leu Gly Ile Lys Val Lys Ser Val Ile Glu Glu Glu Lys Cys Val Ile	
325 330 335	

```

cct atg gga gga cca ctt ccg cgg att cct caa aat gtt atg gct att      1056
Pro Met Gly Gly Pro Leu Pro Arg Ile Pro Gln Asn Val Met Ala Ile
          340                      345                      350

ggg ggg aat tca ggg ata gtt cat cca tca aca ggg tac atg gtg gct      1104
Gly Gly Asn Ser Gly Ile Val His Pro Ser Thr Gly Tyr Met Val Ala
          355                      360                      365

agg agc atg gct tta gca cca gta cta gct gaa gcc atc gtc gag ggg      1152
Arg Ser Met Ala Leu Ala Pro Val Leu Ala Glu Ala Ile Val Glu Gly
          370                      375                      380

ctt ggc tca aca aga atg ata aga ggg tct caa ctt tac cat aga gtt      1200
Leu Gly Ser Thr Arg Met Ile Arg Gly Ser Gln Leu Tyr His Arg Val
          385                      390                      395                      400

tgg aat ggt ttg tgg cct ttg gat aga aga tgt gtt aga gaa tgt tat      1248
Trp Asn Gly Leu Trp Pro Leu Asp Arg Arg Cys Val Arg Glu Cys Tyr
          405                      410                      415

tca ttt ggg atg gag aca ttg ttg aag ctt gat ttg aaa ggg act agg      1296
Ser Phe Gly Met Glu Thr Leu Leu Lys Leu Asp Leu Lys Gly Thr Arg
          420                      425                      430

aga ttg ttt gac gct ttc ttt gat ctt gat cct aaa tac tgg caa ggg      1344
Arg Leu Phe Asp Ala Phe Phe Asp Leu Asp Pro Lys Tyr Trp Gln Gly
          435                      440                      445

ttc ctt tct tca aga ttg tct gtc aaa gaa ctt ggt tta ctc agc ttg      1392
Phe Leu Ser Ser Arg Leu Ser Val Lys Glu Leu Gly Leu Leu Ser Leu
          450                      455                      460

tgt ctt ttc gga cat ggc tca aac atg act agg ttg gat att gtt aca      1440
Cys Leu Phe Gly His Gly Ser Asn Met Thr Arg Leu Asp Ile Val Thr
          465                      470                      475                      480

aaa tgt cct ctt cct ttg gtt aga ctg att ggc aat cta gca ata gag      1488
Lys Cys Pro Leu Pro Leu Val Arg Leu Ile Gly Asn Leu Ala Ile Glu
          485                      490                      495

agc ctt tgaatgtgaa aagtttgaat cattttcttc attttaattt ctttgattat      1544
Ser Leu

tttcatattt tctcaattgc aaaagtgaga taagagctac atactgtcaa caaataaact      1604

actattggaa agttaaaata tgtgtttggt gtatgttatt ctaatggaat ggattttgta      1664

aa                                                                 1666

```

&lt;210&gt; 48

&lt;211&gt; 498

&lt;212&gt; PRT

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 48

Met Glu Ala Leu Leu Lys Pro Phe Pro Ser Leu Leu Leu Ser Ser Pro  
 1 5 10 15

Thr Pro His Arg Ser Ile Phe Gln Gln Asn Pro Ser Phe Leu Ser Pro  
 20 25 30

Thr Thr Lys Lys Lys Ser Arg Lys Cys Leu Leu Arg Asn Lys Ser Ser  
 35 40 45

Lys Leu Phe Cys Ser Phe Leu Asp Leu Ala Pro Thr Ser Lys Pro Glu  
 50 55 60

Ser Leu Asp Val Asn Ile Ser Trp Val Asp Pro Asn Ser Asn Arg Ala  
 65 70 75 80

Gln Phe Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Arg Leu  
 85 90 95

Ala Glu Gln Val Ser Lys Tyr Gly Ile Lys Val Cys Cys Val Asp Pro  
 100 105 110

Ser Pro Leu Ser Met Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu  
 115 120 125

Phe Glu Asn Leu Gly Leu Glu Asn Cys Leu Asp His Lys Trp Pro Met  
 130 135 140

Thr Cys Val His Ile Asn Asp Asn Lys Thr Lys Tyr Leu Gly Arg Pro  
 145 150 155 160

Tyr Gly Arg Val Ser Arg Lys Lys Leu Lys Leu Lys Leu Leu Asn Ser  
 165 170 175

Cys Val Glu Asn Arg Val Lys Phe Tyr Lys Ala Lys Val Trp Lys Val  
 180 185 190

Glu His Glu Glu Phe Glu Ser Ser Ile Val Cys Asp Asp Gly Lys Lys  
 195 200 205

Ile Arg Gly Ser Leu Val Val Asp Ala Ser Gly Phe Ala Ser Asp Phe  
 210 215 220

Ile Glu Tyr Asp Arg Pro Arg Asn His Gly Tyr Gln Ile Ala His Gly  
 225 230 235 240

Val Leu Val Glu Val Asp Asn His Pro Phe Asp Leu Asp Lys Met Val  
 245 250 255

Leu Met Asp Trp Arg Asp Ser His Leu Gly Asn Glu Pro Tyr Leu Arg  
 260 265 270

Val Asn Asn Ala Lys Glu Pro Thr Phe Leu Tyr Ala Met Pro Phe Asp  
 275 280 285

Arg Asp Leu Val Phe Leu Glu Glu Thr Ser Leu Val Ser Arg Pro Val  
 290 295 300

Leu Ser Tyr Met Glu Val Lys Arg Arg Met Val Ala Arg Leu Arg His  
 305 310 315 320

Leu Gly Ile Lys Val Lys Ser Val Ile Glu Glu Glu Lys Cys Val Ile  
 325 330 335

Pro Met Gly Gly Pro Leu Pro Arg Ile Pro Gln Asn Val Met Ala Ile  
 340 345 350

Gly Gly Asn Ser Gly Ile Val His Pro Ser Thr Gly Tyr Met Val Ala  
 355 360 365

Arg Ser Met Ala Leu Ala Pro Val Leu Ala Glu Ala Ile Val Glu Gly  
 370 375 380

Leu Gly Ser Thr Arg Met Ile Arg Gly Ser Gln Leu Tyr His Arg Val  
 385 390 395 400

Trp Asn Gly Leu Trp Pro Leu Asp Arg Arg Cys Val Arg Glu Cys Tyr  
 405 410 415

Ser Phe Gly Met Glu Thr Leu Leu Lys Leu Asp Leu Lys Gly Thr Arg  
 420 425 430

Arg Leu Phe Asp Ala Phe Phe Asp Leu Asp Pro Lys Tyr Trp Gln Gly  
 435 440 445

Phe Leu Ser Ser Arg Leu Ser Val Lys Glu Leu Gly Leu Leu Ser Leu  
 450 455 460

Cys Leu Phe Gly His Gly Ser Asn Met Thr Arg Leu Asp Ile Val Thr  
 465 470 475 480

Lys Cys Pro Leu Pro Leu Val Arg Leu Ile Gly Asn Leu Ala Ile Glu  
 485 490 495

Ser Leu

<210> 49

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3)..(971)

<223>

<400> 49

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47  
 Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile  
 1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95  
 Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu  
 20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143  
 Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala  
 35 40 45

cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg	191
Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser	
50 55 60	
tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga	239
Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly	
65 70 75	
acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca	287
Thr Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala	
80 85 90 95	
ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa	335
Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys	
100 105 110	
cgg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc	383
Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly	
115 120 125	
gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac	431
Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His	
130 135 140	
atg acc gtg ggc ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc	479
Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu	
145 150 155	
ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat	527
Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr	
160 165 170 175	
gca cac aaa gcc atc tgg cat gag tcg cct ctg ggc tgg ctg ctg cac	575
Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His	
180 185 190	
aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg	623
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu	
195 200 205	
ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc	671
Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly	
210 215 220	
ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg	719
Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu	
225 230 235	
ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg	767
Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu	
240 245 250 255	
gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg	815
Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met	
260 265 270	

aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt 863  
 Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly  
 275 280 285  
 ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att 911  
 Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile  
 290 295 300  
 cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg 959  
 Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp  
 305 310 315  
 tcc aag cgg tag ggtgcggaac caggcacgct ggtttcacac ctcatgcctg 1011  
 Ser Lys Arg  
 320  
 tgataagggtg tggctagagc gatgcgtgtg agacgggtat gtcacgggtcg actgggtctga 1071  
 tggccaatgg catcgggccat gtctgggtcat cacgggctgg ttgcctgggt gaagggtgatg 1131  
 cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191  
 caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgcc 1251  
 catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1311  
 gtgcagcaaa ctatattcac ctagggtgtg tggtaggatc aggtgaggcc ttgcacattg 1371  
 catgatgtac tcgtcatggt gtgttggtga gaggatggat gtggatggat gtgtattctc 1431  
 agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattcct tgagagggga 1491  
 ggctcgtgcc agaaatgggtg agtggatgac tgtgacgctg tacattgcag gcagggtgaga 1551  
 tgcactgtct cgattgtaaa atacattcag atgcacaaaaa aaaaaaaaaa aaaaaaa 1608

&lt;210&gt; 50

&lt;211&gt; 322

&lt;212&gt; PRT

&lt;213&gt; Haematococcus pluvialis

&lt;400&gt; 50

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly  
 1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser  
 20 25 30



Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg  
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu  
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr  
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu  
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg  
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val  
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met  
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu  
145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala  
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys  
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe  
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe  
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly  
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val  
245 250 255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys  
 260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly  
 275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro  
 290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser  
 305 310 315 320

Lys Arg

<210> 51

<211> 1163

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(942)

<223>

<400> 51

att cgg cac gag att tca gcc tcc gct agt tcc cga acc att cgc ctc 48  
 Ile Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu  
 1 5 10 15

cgt cat aac ccg ttt ctc agt cca aaa tcc gcc tca acc gcc ccg ccg 96  
 Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro  
 20 25 30

gtt ctg ttc ttc tct ccg tta act cgc aat ttt ggc gca att ttg ctg 144  
 Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu  
 35 40 45

tct aga aga aag ccg aga ttg gcg gtt tgt ttt gtg ctg gag aat gag 192

Ser	Arg	Arg	Lys	Pro	Arg	Leu	Ala	Val	Cys	Phe	Val	Leu	Glu	Asn	Glu	
50						55					60					
aaa	ttg	aat	agt	act	atc	gaa	agt	gag	agt	gaa	gta	ata	gag	gat	cgg	240
Lys	Leu	Asn	Ser	Thr	Ile	Glu	Ser	Glu	Ser	Glu	Val	Ile	Glu	Asp	Arg	
65					70					75					80	
ata	caa	gta	gag	att	aat	gag	gag	aag	agt	tta	gct	gcc	agt	tgg	ctg	288
Ile	Gln	Val	Glu	Ile	Asn	Glu	Glu	Lys	Ser	Leu	Ala	Ala	Ser	Trp	Leu	
				85					90					95		
gcg	gag	aaa	ttg	gcg	agg	aag	aaa	tcg	gag	agg	ttt	act	tat	ctt	gtg	336
Ala	Glu	Lys	Leu	Ala	Arg	Lys	Lys	Ser	Glu	Arg	Phe	Thr	Tyr	Leu	Val	
			100					105						110		
gca	gct	gtg	atg	tct	agt	ttg	ggg	att	act	tct	atg	gcg	att	ttg	gcg	384
Ala	Ala	Val	Met	Ser	Ser	Leu	Gly	Ile	Thr	Ser	Met	Ala	Ile	Leu	Ala	
			115				120					125				
gtt	tat	tac	aga	ttt	tca	tgg	caa	atg	gag	ggt	gga	gaa	gtg	cct	ttt	432
Val	Tyr	Arg	Phe	Ser	Trp	Gln	Met	Glu	Gly	Gly	Glu	Val	Pro	Phe		
	130					135					140					
tct	gaa	atg	tta	gct	aca	ttc	act	ctc	tcg	ttt	ggc	gct	gcc	gta	gga	480
Ser	Glu	Met	Leu	Ala	Thr	Phe	Thr	Leu	Ser	Phe	Gly	Ala	Ala	Val	Gly	
145					150					155					160	
atg	gag	tac	tgg	gcg	aga	tgg	gct	cat	aga	gca	cta	tgg	cat	gct	tct	528
Met	Glu	Tyr	Trp	Ala	Arg	Trp	Ala	His	Arg	Ala	Leu	Trp	His	Ala	Ser	
				165					170					175		
tta	tgg	cac	atg	cac	gag	tcg	cac	cat	aga	cca	aga	gaa	gga	cct	ttt	576
Leu	Trp	His	Met	His	Glu	Ser	His	His	Arg	Pro	Arg	Glu	Gly	Pro	Phe	
			180					185					190			
gag	atg	aac	gac	gtt	ttc	gcc	ata	aca	aat	gct	gtt	cca	gct	ata	ggc	624
Glu	Met	Asn	Asp	Val	Phe	Ala	Ile	Thr	Asn	Ala	Val	Pro	Ala	Ile	Gly	
		195					200					205				
ctt	ctt	tcc	tac	ggc	ttc	ttc	cat	aaa	ggg	atc	gtc	cct	ggc	ctc	tgt	672
Leu	Leu	Ser	Tyr	Gly	Phe	Phe	His	Lys	Gly	Ile	Val	Pro	Gly	Leu	Cys	
	210					215					220					
ttc	ggc	gct	gga	ttg	ggg	atc	aca	gta	ttt	ggg	atg	gct	tac	atg	ttc	720
Phe	Gly	Ala	Gly	Leu	Gly	Ile	Thr	Val	Phe	Gly	Met	Ala	Tyr	Met	Phe	
225					230					235					240	
gtt	cac	gat	gga	ctg	gtt	cat	aag	aga	ttt	ccc	gta	ggg	cct	att	gcc	768
Val	His	Asp	Gly	Leu	Val	His	Lys	Arg	Phe	Pro	Val	Gly	Pro	Ile	Ala	
				245					250					255		
aac	gtg	cct	tac	ttt	cgg	agg	gta	gct	gca	gca	cat	cag	ctt	cat	cac	816
Asn	Val	Pro	Tyr	Phe	Arg	Arg	Val	Ala	Ala	Ala	His	Gln	Leu	His	His	
			260					265					270			
tcg	gac	aaa	ttt	gat	ggc	gtc	cca	tat	ggc	ttg	ttt	cta	gga	cct	aag	864

Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys  
 275 280 285

gaa ttg gaa gaa gta gga gga ctt gaa gag tta gaa aag gaa gtc aac 912  
 Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Val Asn  
 290 295 300

cga agg att aaa att tct aag gga tta tta tgatcaaaag atacgtctga 962  
 Arg Arg Ile Lys Ile Ser Lys Gly Leu Leu  
 305 310

taataataaa atgcgattgt atttaggctg tagattatta ttgggaaaaa gatagaaaga 1022

tatatatatg aatataatat aaaatgcaac aagctttcta tggagaagac cttttctttt 1082

ttggtacctg tacgtaaaag gtgaacaatt tgatgtccta gtacttgttg acaaaccaga 1142

agaacgataa ttcaaaacaa a 1163

&lt;210&gt; 52

&lt;211&gt; 314

&lt;212&gt; PRT

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 52

Ile Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu  
 1 5 10 15

Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro  
 20 25 30

Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu  
 35 40 45

Ser Arg Arg Lys Pro Arg Leu Ala Val Cys Phe Val Leu Glu Asn Glu  
 50 55 60

Lys Leu Asn Ser Thr Ile Glu Ser Glu Ser Glu Val Ile Glu Asp Arg  
 65 70 75 80

Ile Gln Val Glu Ile Asn Glu Glu Lys Ser Leu Ala Ala Ser Trp Leu  
 85 90 95

Ala Glu Lys Leu Ala Arg Lys Lys Ser Glu Arg Phe Thr Tyr Leu Val  
100 105 110

Ala Ala Val Met Ser Ser Leu Gly Ile Thr Ser Met Ala Ile Leu Ala  
115 120 125

Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly Glu Val Pro Phe  
130 135 140

Ser Glu Met Leu Ala Thr Phe Thr Leu Ser Phe Gly Ala Ala Val Gly  
145 150 155 160

Met Glu Tyr Trp Ala Arg Trp Ala His Arg Ala Leu Trp His Ala Ser  
165 170 175

Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe  
180 185 190

Glu Met Asn Asp Val Phe Ala Ile Thr Asn Ala Val Pro Ala Ile Gly  
195 200 205

Leu Leu Ser Tyr Gly Phe Phe His Lys Gly Ile Val Pro Gly Leu Cys  
210 215 220

Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe  
225 230 235 240

Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala  
245 250 255

Asn Val Pro Tyr Phe Arg Arg Val Ala Ala Ala His Gln Leu His His  
260 265 270

Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys  
275 280 285

Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Val Asn  
290 295 300

Arg Arg Ile Lys Ile Ser Lys Gly Leu Leu  
305 310

&lt;210&gt; 53

&lt;211&gt; 1779

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (1779)

&lt;223&gt;

&lt;400&gt; 53

atg	gat	ctc	cgt	cgg	agg	cct	cct	aaa	cca	ccg	gtt	acc	aac	aac	aac	48
Met	Asp	Leu	Arg	Arg	Arg	Pro	Pro	Lys	Pro	Pro	Val	Thr	Asn	Asn	Asn	
1			5					10					15			

aac	tcc	aac	gga	tct	ttc	cgt	tct	tat	cag	cct	cgc	act	tcc	gat	gac	96
Asn	Ser	Asn	Gly	Ser	Phe	Arg	Ser	Tyr	Gln	Pro	Arg	Thr	Ser	Asp	Asp	
			20					25					30			

gat	cat	cgt	cgc	cgg	gct	aca	aca	att	gct	cct	cca	ccg	aaa	gca	tcc	144
Asp	His	Arg	Arg	Arg	Ala	Thr	Thr	Ile	Ala	Pro	Pro	Pro	Lys	Ala	Ser	
			35					40					45			

gac	gcg	ctt	cct	ctt	ccg	tta	tat	ctc	aca	aac	gcc	gtt	ttc	ttc	acg	192
Asp	Ala	Leu	Pro	Leu	Pro	Leu	Tyr	Leu	Thr	Asn	Ala	Val	Phe	Phe	Thr	
			50				55				60					

ctc	ttc	ttc	tcc	gtc	gcg	tat	tac	ctc	ctc	cac	cgg	tgg	cgt	gac	aag	240
Leu	Phe	Phe	Ser	Val	Ala	Tyr	Tyr	Leu	Leu	His	Arg	Trp	Arg	Asp	Lys	
65				70						75				80		

atc	cgt	tac	aat	acg	cct	ctt	cac	gtc	gtc	act	atc	aca	gaa	ctc	ggc	288
Ile	Arg	Tyr	Asn	Thr	Pro	Leu	His	Val	Val	Thr	Ile	Thr	Glu	Leu	Gly	
				85					90					95		

gcc	att	att	gct	ctc	atc	gct	tcg	ttt	atc	tat	ctc	cta	ggg	ttt	ttt	336
Ala	Ile	Ile	Ala	Leu	Ile	Ala	Ser	Phe	Ile	Tyr	Leu	Leu	Gly	Phe	Phe	
			100					105					110			

ggc	att	gac	ttt	gtt	cag	tca	ttt	atc	tca	cgt	gcc	tct	ggc	gat	gct	384
Gly	Ile	Asp	Phe	Val	Gln	Ser	Phe	Ile	Ser	Arg	Ala	Ser	Gly	Asp	Ala	
			115				120					125				

tgg	gat	ctc	gcc	gat	acg	atc	gat	gat	gat	gac	cac	cgc	ctt	gtc	acg	432
Trp	Asp	Leu	Ala	Asp	Thr	Ile	Asp	Asp	Asp	Asp	His	Arg	Leu	Val	Thr	
			130				135					140				

tgc tct cca ccg act ccg atc gtt tcc gtt gct aaa tta cct aat ccg	480
Cys Ser Pro Pro Thr Pro Ile Val Ser Val Ala Lys Leu Pro Asn Pro	
145 150 155 160	
gaa cct att gtt acc gaa tcg ctt cct gag gaa gac gag gag att gtg	528
Glu Pro Ile Val Thr Glu Ser Leu Pro Glu Glu Asp Glu Glu Ile Val	
165 170 175	
aaa tcg gtt atc gac gga gtt att cca tcg tac tcg ctt gaa tct cgt	576
Lys Ser Val Ile Asp Gly Val Ile Pro Ser Tyr Ser Leu Glu Ser Arg	
180 185 190	
ctc ggt gat tgc aaa aga gcg gcg tcg att cgt cgt gag gcg ttg cag	624
Leu Gly Asp Cys Lys Arg Ala Ala Ser Ile Arg Arg Glu Ala Leu Gln	
195 200 205	
aga gtc acc ggg aga tcg att gaa ggg tta ccg ttg gat gga ttt gat	672
Arg Val Thr Gly Arg Ser Ile Glu Gly Leu Pro Leu Asp Gly Phe Asp	
210 215 220	
tat gaa tcg att ttg ggg caa tgc tgt gag atg cct gtt gga tac att	720
Tyr Glu Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Val Gly Tyr Ile	
225 230 235 240	
cag att cct gtt ggg att gct ggt cca ttg ttg ctt gat ggt tat gag	768
Gln Ile Pro Val Gly Ile Ala Gly Pro Leu Leu Leu Asp Gly Tyr Glu	
245 250 255	
tac tct gtt cct atg gct aca acc gaa ggt tgt ttg gtt gct agc act	816
Tyr Ser Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr	
260 265 270	
aac aga ggc tgc aag gct atg ttt atc tct ggt ggc gcc acc agt acc	864
Asn Arg Gly Cys Lys Ala Met Phe Ile Ser Gly Gly Ala Thr Ser Thr	
275 280 285	
gtt ctt aag gac ggt atg acc cga gca cct gtt gtt cgg ttc gct tcg	912
Val Leu Lys Asp Gly Met Thr Arg Ala Pro Val Val Arg Phe Ala Ser	
290 295 300	
gcg aga cga gct tcg gag ctt aag ttt ttc ttg gag aat cca gag aac	960
Ala Arg Arg Ala Ser Glu Leu Lys Phe Phe Leu Glu Asn Pro Glu Asn	
305 310 315 320	
ttt gat act ttg gca gta gtc ttc aac agg tcg agt aga ttt gca aga	1008
Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg	
325 330 335	
ctg caa agt gtt aaa tgc aca atc gcg ggg aag aat gct tat gta agg	1056
Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg	
340 345 350	
ttc tgt tgt agt act ggt gat gct atg ggg atg aat atg gtt tct aaa	1104
Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys	
355 360 365	

ggt gtg cag aat gtt ctt gag tat ctt acc gat gat ttc cct gac atg	1152
Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met	
370 375 380	
gat gtg att gga atc tct ggt aac ttc tgt tgc gac aag aaa cct gct	1200
Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala	
385 390 395 400	
gct gtg aac tgg att gag gga cgt ggt aaa tca gtt gtt tgc gag gct	1248
Ala Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala	
405 410 415	
gta atc aga gga gag atc gtg aac aag gtc ttg aaa acg agc gtg gct	1296
Val Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala	
420 425 430	
gct tta gtc gag ctc aac atg ctc aag aac cta gct ggc tct gct gtt	1344
Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val	
435 440 445	
gca ggc tct cta ggt gga ttc aac gct cat gcc agt aac ata gtg tct	1392
Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser	
450 455 460	
gct gta ttc ata gct act ggc caa gat cca gct caa aac gtg gag agt	1440
Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser	
465 470 475 480	
tct caa tgc atc acc atg atg gaa gct att aat gac ggc aaa gat atc	1488
Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile	
485 490 495	
cat atc tca gtc act atg cca tct atc gag gtg ggg aca gtg gga gga	1536
His Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly	
500 505 510	
gga aca cag ctt gca tct caa tca gcg tgt tta aac ctg ctc gga gtt	1584
Gly Thr Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val	
515 520 525	
aaa gga gca agc aca gag tgc ccg gga atg aac gca agg agg cta gcg	1632
Lys Gly Ala Ser Thr Glu Ser Pro Gly Met Asn Ala Arg Arg Leu Ala	
530 535 540	
acg atc gta gcc gga gca gtt tta gct gga gag tta tct tta atg tca	1680
Thr Ile Val Ala Gly Ala Val Leu Ala Gly Glu Leu Ser Leu Met Ser	
545 550 555 560	
gca att gca gct gga cag ctt gtg aga agt cac atg aaa tac aat aga	1728
Ala Ile Ala Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg	
565 570 575	
tcc agc cga gac atc tct gga gca acg aca acg aca aca aca aca	1776
Ser Ser Arg Asp Ile Ser Gly Ala Thr Thr Thr Thr Thr Thr Thr	
580 585 590	



tga

1779

&lt;210&gt; 54

&lt;211&gt; 592

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 54

Met Asp Leu Arg Arg Arg Pro Pro Lys Pro Pro Val Thr Asn Asn Asn  
 1 5 10 15

Asn Ser Asn Gly Ser Phe Arg Ser Tyr Gln Pro Arg Thr Ser Asp Asp  
 20 25 30

Asp His Arg Arg Arg Ala Thr Thr Ile Ala Pro Pro Pro Lys Ala Ser  
 35 40 45

Asp Ala Leu Pro Leu Pro Leu Tyr Leu Thr Asn Ala Val Phe Phe Thr  
 50 55 60

Leu Phe Phe Ser Val Ala Tyr Tyr Leu Leu His Arg Trp Arg Asp Lys  
 65 70 75 80

Ile Arg Tyr Asn Thr Pro Leu His Val Val Thr Ile Thr Glu Leu Gly  
 85 90 95

Ala Ile Ile Ala Leu Ile Ala Ser Phe Ile Tyr Leu Leu Gly Phe Phe  
 100 105 110

Gly Ile Asp Phe Val Gln Ser Phe Ile Ser Arg Ala Ser Gly Asp Ala  
 115 120 125

Trp Asp Leu Ala Asp Thr Ile Asp Asp Asp Asp His Arg Leu Val Thr  
 130 135 140

Cys Ser Pro Pro Thr Pro Ile Val Ser Val Ala Lys Leu Pro Asn Pro  
 145 150 155 160

Glu Pro Ile Val Thr Glu Ser Leu Pro Glu Glu Asp Glu Glu Ile Val  
165 170 175

Lys Ser Val Ile Asp Gly Val Ile Pro Ser Tyr Ser Leu Glu Ser Arg  
180 185 190

Leu Gly Asp Cys Lys Arg Ala Ala Ser Ile Arg Arg Glu Ala Leu Gln  
195 200 205

Arg Val Thr Gly Arg Ser Ile Glu Gly Leu Pro Leu Asp Gly Phe Asp  
210 215 220

Tyr Glu Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Val Gly Tyr Ile  
225 230 235 240

Gln Ile Pro Val Gly Ile Ala Gly Pro Leu Leu Leu Asp Gly Tyr Glu  
245 250 255

Tyr Ser Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr  
260 265 270

Asn Arg Gly Cys Lys Ala Met Phe Ile Ser Gly Gly Ala Thr Ser Thr  
275 280 285

Val Leu Lys Asp Gly Met Thr Arg Ala Pro Val Val Arg Phe Ala Ser  
290 295 300

Ala Arg Arg Ala Ser Glu Leu Lys Phe Phe Leu Glu Asn Pro Glu Asn  
305 310 315 320

Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg  
325 330 335

Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg  
340 345 350

Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys  
355 360 365

Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met  
370 375 380

Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala  
385 390 395 400

Ala Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala  
405 410 415

Val Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala  
420 425 430

Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val  
435 440 445

Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser  
450 455 460

Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser  
465 470 475 480

Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile  
485 490 495

His Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly  
500 505 510

Gly Thr Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val  
515 520 525

Lys Gly Ala Ser Thr Glu Ser Pro Gly Met Asn Ala Arg Arg Leu Ala  
530 535 540

Thr Ile Val Ala Gly Ala Val Leu Ala Gly Glu Leu Ser Leu Met Ser  
545 550 555 560

Ala Ile Ala Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg  
565 570 575

Ser Ser Arg Asp Ile Ser Gly Ala Thr Thr Thr Thr Thr Thr Thr  
580 585 590

<210> 55

<211> 1401

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana ISPH

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1401)

&lt;223&gt;

&lt;400&gt; 55

atg gct gtt gcg ctc caa ttc agc cga tta tgc gtt cga ccg gat act	48
Met Ala Val Ala Leu Gln Phe Ser Arg Leu Cys Val Arg Pro Asp Thr	
1                      5                      10                      15	

ttc gtg cgg gag aat cat ctc tct gga tcc gga tct ctc cgc cgc cgg	96
Phe Val Arg Glu Asn His Leu Ser Gly Ser Gly Ser Leu Arg Arg Arg	
20                      25                      30	

aaa gct tta tca gtc cgg tgc tcg tct ggc gat gag aac gct cct tcg	144
Lys Ala Leu Ser Val Arg Cys Ser Ser Gly Asp Glu Asn Ala Pro Ser	
35                      40                      45	

cca tcg gtg gtg atg gac tcc gat ttc gac gcc aag gtg ttc cgt aag	192
Pro Ser Val Val Met Asp Ser Asp Phe Asp Ala Lys Val Phe Arg Lys	
50                      55                      60	

aac ttg acg aga agc gat aat tac aat cgt aaa ggg ttc ggt cat aag	240
Asn Leu Thr Arg Ser Asp Asn Tyr Asn Arg Lys Gly Phe Gly His Lys	
65                      70                      75                      80	

gag gag aca ctc aag ctc atg aat cga gag tac acc agt gat ata ttg	288
Glu Glu Thr Leu Lys Leu Met Asn Arg Glu Tyr Thr Ser Asp Ile Leu	
85                      90                      95	

gag aca ctg aaa aca aat ggg tat act tat tct tgg gga gat gtt act	336
Glu Thr Leu Lys Thr Asn Gly Tyr Thr Tyr Ser Trp Gly Asp Val Thr	
100                      105                      110	

gtg aaa ctc gct aaa gca tat ggt ttt tgc tgg ggt gtt gag cgt gct	384
Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala	
115                      120                      125	

gtt cag att gca tat gaa gca cga aag cag ttt cca gag gag agg ctt	432
Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu	
130                      135                      140	

tgg att act aac gaa atc att cat aac ccg acc gtc aat aag agg ttg	480
Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu	
145                      150                      155                      160	

gaa gat atg gat gtt aaa att att ccg gtt gag gat tca aag aaa cag	528
Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln	
165 170 175	
ttt gat gta gta gag aaa gat gat gtg gtt atc ctt cct gcg ttt gga	576
Phe Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly	
180 185 190	
gct ggt gtt gac gag atg tat gtt ctt aat gat aaa aag gtg caa att	624
Ala Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile	
195 200 205	
gtt gac acg act tgt cct tgg gtg aca aag gtc tgg aac acg gtt gag	672
Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu	
210 215 220	
aag cac aag aag ggg gaa tac aca tca gta atc cat ggt aaa tat aat	720
Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn	
225 230 235 240	
cat gaa gag acg att gca act gcg tct ttt gca gga aag tac atc att	768
His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile	
245 250 255	
gta aag aac atg aaa gag gca aat tac gtt tgt gat tac att ctc ggt	816
Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly	
260 265 270	
ggc caa tac gat gga tct agc tcc aca aaa gag gag ttc atg gag aaa	864
Gly Gln Tyr Asp Gly Ser Ser Ser Thr Lys Glu Glu Phe Met Glu Lys	
275 280 285	
ttc aaa tac gca att tcg aag ggt ttc gat ccc gac aat gac ctt gtc	912
Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val	
290 295 300	
aaa gtt ggt att gca aac caa aca acg atg cta aag gga gaa aca gag	960
Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu	
305 310 315 320	
gag ata gga aga tta ctc gag aca aca atg atg cgc aag tat gga gtg	1008
Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val	
325 330 335	
gaa aat gta agc gga cat ttc atc agc ttc aac aca ata tgc gac gct	1056
Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala	
340 345 350	
act caa gag cga caa gac gca atc tat gag cta gtg gaa gag aag att	1104
Thr Gln Glu Arg Gln Asp Ala Ile Tyr Glu Leu Val Glu Glu Lys Ile	
355 360 365	
gac ctc atg cta gtg gtt ggc gga tgg aat tca agt aac acc tct cac	1152
Asp Leu Met Leu Val Val Gly Gly Trp Asn Ser Ser Asn Thr Ser His	
370 375 380	

ctt cag gaa atc tca gag gca cgg gga atc cca tct tac tgg atc gat 1200  
 Leu Gln Glu Ile Ser Glu Ala Arg Gly Ile Pro Ser Tyr Trp Ile Asp  
 385 390 395 400

agt gag aaa cgg ata gga cct ggg aat aaa ata gcc tat aag ctc cac 1248  
 Ser Glu Lys Arg Ile Gly Pro Gly Asn Lys Ile Ala Tyr Lys Leu His  
 405 410 415

tat gga gaa ctg gtc gag aag gaa aac ttt ctc cca aag gga cca ata 1296  
 Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile  
 420 425 430

aca atc ggt gtg aca tca ggt gca tca acc ccg gat aag gtc gtg gaa 1344  
 Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu  
 435 440 445

gat gct ttg gtg aag gtg ttc gac att aaa cgt gaa gag tta ttg cag 1392  
 Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln  
 450 455 460

ctg gct tga 1401  
 Leu Ala  
 465

<210> 56

<211> 466

<212> PRT

<213> Arabidopsis thaliana ISPH

<400> 56

Met Ala Val Ala Leu Gln Phe Ser Arg Leu Cys Val Arg Pro Asp Thr  
 1 5 10 15

Phe Val Arg Glu Asn His Leu Ser Gly Ser Gly Ser Leu Arg Arg Arg  
 20 25 30

Lys Ala Leu Ser Val Arg Cys Ser Ser Gly Asp Glu Asn Ala Pro Ser  
 35 40 45

Pro Ser Val Val Met Asp Ser Asp Phe Asp Ala Lys Val Phe Arg Lys  
 50 55 60

Asn Leu Thr Arg Ser Asp Asn Tyr Asn Arg Lys Gly Phe Gly His Lys  
 65 70 75 80

Glu Glu Thr Leu Lys Leu Met Asn Arg Glu Tyr Thr Ser Asp Ile Leu  
85 90 95

Glu Thr Leu Lys Thr Asn Gly Tyr Thr Tyr Ser Trp Gly Asp Val Thr  
100 105 110

Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala  
115 120 125

Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu  
130 135 140

Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu  
145 150 155 160

Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln  
165 170 175

Phe Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly  
180 185 190

Ala Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile  
195 200 205

Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu  
210 215 220

Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn  
225 230 235 240

His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile  
245 250 255

Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly  
260 265 270

Gly Gln Tyr Asp Gly Ser Ser Ser Thr Lys Glu Glu Phe Met Glu Lys  
275 280 285

Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val  
290 295 300

Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu  
 305 310 315 320

Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val  
 325 330 335

Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala  
 340 345 350

Thr Gln Glu Arg Gln Asp Ala Ile Tyr Glu Leu Val Glu Glu Lys Ile  
 355 360 365

Asp Leu Met Leu Val Val Gly Gly Trp Asn Ser Ser Asn Thr Ser His  
 370 375 380

Leu Gln Glu Ile Ser Glu Ala Arg Gly Ile Pro Ser Tyr Trp Ile Asp  
 385 390 395 400

Ser Glu Lys Arg Ile Gly Pro Gly Asn Lys Ile Ala Tyr Lys Leu His  
 405 410 415

Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile  
 420 425 430

Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu  
 435 440 445

Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln  
 450 455 460

Leu Ala  
 465

<210> 57

<211> 2160

<212> DNA

<213> Lycopersicon esculentum



&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(2160)

&lt;223&gt;

&lt;400&gt; 57

atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt	48
Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly	
1 5 10 15	
gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att	96
Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile	
20 25 30	
cat gga aca gat ctg cag ttt ttg ttc caa cac aag ctt act cat gag	144
His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu	
35 40 45	
gtc aag aaa agg tca cgt gtg gtt cag gct tcc tta tca gaa tct gga	192
Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly	
50 55 60	
gaa tac tac aca cag aga ccg cca acg cct att ttg gac act gtg aac	240
Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn	
65 70 75 80	
tat ccc att cat atg aaa aat ctg tct ctg aag gaa ctt aaa caa cta	288
Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu	
85 90 95	
gca gat gaa cta agg tca gat aca att ttc aat gta tca aag act ggg	336
Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly	
100 105 110	
ggg cac ctt ggc tca agt ctt ggt gtt gtt gag ctg act gtt gct ctt	384
Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu	
115 120 125	
cat tat gtc ttc aat gca ccg caa gat agg att ctc tgg gat gtt ggt	432
His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly	
130 135 140	
cat cag tct tat cct cac aaa atc ttg act ggt aga agg gac aag atg	480
His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met	
145 150 155 160	
tcg aca tta agg cag aca gat ggt ctt gca gga ttt act aag cga tcg	528
Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser	
165 170 175	
gag agt gaa tat gat tgc ttt ggc acc ggc cac agt tcc acc acc atc	576

Glu Ser Glu Tyr Asp Cys Phe Gly Thr Gly His Ser Ser Thr Thr Ile	
180 185 190	
tca gca ggc cta ggg atg gct gtt ggt aga gat cta aaa gga aga aac	624
Ser Ala Gly Leu Gly Met Ala Val Gly Arg Asp Leu Lys Gly Arg Asn	
195 200 205	
aac aat gtt att gcc gta ata ggt gat ggt gcc atg aca gca ggt caa	672
Asn Asn Val Ile Ala Val Ile Gly Asp Gly Ala Met Thr Ala Gly Gln	
210 215 220	
gct tat gaa gcc atg aat aat gct ggt tac ctg gac tct gac atg att	720
Ala Tyr Glu Ala Met Asn Asn Ala Gly Tyr Leu Asp Ser Asp Met Ile	
225 230 235 240	
gtt atc tta aac gac aat aga caa gtt tct tta cct act gct act ctg	768
Val Ile Leu Asn Asp Asn Arg Gln Val Ser Leu Pro Thr Ala Thr Leu	
245 250 255	
gat ggg cca gtt gct cct gtt gga gct cta agt agt gct ttg agc agg	816
Asp Gly Pro Val Ala Pro Val Gly Ala Leu Ser Ser Ala Leu Ser Arg	
260 265 270	
tta cag tct aat agg cct ctc aga gaa cta aga gaa gtc gca aag gga	864
Leu Gln Ser Asn Arg Pro Leu Arg Glu Leu Arg Glu Val Ala Lys Gly	
275 280 285	
gtt act aag cag att ggt ggt cct atg cat gag ctt gct gca aaa gtt	912
Val Thr Lys Gln Ile Gly Gly Pro Met His Glu Leu Ala Ala Lys Val	
290 295 300	
gat gaa tat gct cgt ggc atg att agt ggt tct gga tca aca ttg ttt	960
Asp Glu Tyr Ala Arg Gly Met Ile Ser Gly Ser Gly Ser Thr Leu Phe	
305 310 315 320	
gaa gaa ctt gga ctt tac tat att ggt cct gtg gat ggt cac aac att	1008
Glu Glu Leu Gly Leu Tyr Tyr Ile Gly Pro Val Asp Gly His Asn Ile	
325 330 335	
gat gat cta att gcg att ctc aaa gag gtt aga agt act aaa aca aca	1056
Asp Asp Leu Ile Ala Ile Leu Lys Glu Val Arg Ser Thr Lys Thr Thr	
340 345 350	
ggt cca gta ctg atc cat gtt gtc act gag aaa ggc aga ggt tat cca	1104
Gly Pro Val Leu Ile His Val Val Thr Glu Lys Gly Arg Gly Tyr Pro	
355 360 365	
tat gct gag aga gct gca gat aag tat cat gga gtt gcc aag ttt gat	1152
Tyr Ala Glu Arg Ala Ala Asp Lys Tyr His Gly Val Ala Lys Phe Asp	
370 375 380	
cca gca aca gga aag caa ttc aaa gcc agt gcc aag aca cag tcc tat	1200
Pro Ala Thr Gly Lys Gln Phe Lys Ala Ser Ala Lys Thr Gln Ser Tyr	
385 390 395 400	
aca aca tat ttt gcc gag gct tta att gca gaa gca gaa gca gat aaa	1248

Thr Thr Tyr Phe Ala Glu Ala Leu Ile Ala Glu Ala Glu Ala Asp Lys	
405 410 415	
gac att gtt gca atc cat gct gcc atg ggg ggt ggg acc gga atg aac	1296
Asp Ile Val Ala Ile His Ala Ala Met Gly Gly Gly Thr Gly Met Asn	
420 425 430	
ctt ttc cat cgt cgc ttc cca aca agg tgt ttt gat gtt gga ata gca	1344
Leu Phe His Arg Arg Phe Pro Thr Arg Cys Phe Asp Val Gly Ile Ala	
435 440 445	
gaa caa cat gca gta acc ttt gct gct gga ttg gct tgt gaa ggc att	1392
Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Cys Glu Gly Ile	
450 455 460	
aaa cct ttc tgt gca atc tat tcg tct ttc atg cag agg gct tat gac	1440
Lys Pro Phe Cys Ala Ile Tyr Ser Ser Phe Met Gln Arg Ala Tyr Asp	
465 470 475 480	
cag gta gtg cat gac gtt gat ttg caa aag ctg ccc gtg agg ttt gca	1488
Gln Val Val His Asp Val Asp Leu Gln Lys Leu Pro Val Arg Phe Ala	
485 490 495	
atg gac aga gca ggt ctt gtt gga gca gat ggt cca aca cat tgt ggt	1536
Met Asp Arg Ala Gly Leu Val Gly Ala Asp Gly Pro Thr His Cys Gly	
500 505 510	
gca ttt gat gtt act tac atg gca tgt ctt cct aac atg gtt gta atg	1584
Ala Phe Asp Val Thr Tyr Met Ala Cys Leu Pro Asn Met Val Val Met	
515 520 525	
gct cct tct gat gaa gcg gag cta ttt cac atg gta gca act gct gcc	1632
Ala Pro Ser Asp Glu Ala Glu Leu Phe His Met Val Ala Thr Ala Ala	
530 535 540	
gcc att gat gac aga cca agt tgt ttt aga tac cca aga gga aat ggg	1680
Ala Ile Asp Asp Arg Pro Ser Cys Phe Arg Tyr Pro Arg Gly Asn Gly	
545 550 555 560	
atc ggt gta gag ctt ccg gct gga aac aaa gga att cct ctt gag gtt	1728
Ile Gly Val Glu Leu Pro Ala Gly Asn Lys Gly Ile Pro Leu Glu Val	
565 570 575	
ggt aaa ggt agg ata ttg att gag ggg gag aga gtg gct cta ttg gga	1776
Gly Lys Gly Arg Ile Leu Ile Glu Gly Glu Arg Val Ala Leu Leu Gly	
580 585 590	
tat ggc tca gca gtg cag aac tgt ttg gat gct gct att gtg cta gaa	1824
Tyr Gly Ser Ala Val Gln Asn Cys Leu Asp Ala Ala Ile Val Leu Glu	
595 600 605	
tcc cgc ggc tta caa gta aca gtt gca gat gca cgt ttc tgc aaa cca	1872
Ser Arg Gly Leu Gln Val Thr Val Ala Asp Ala Arg Phe Cys Lys Pro	
610 615 620	
ctg gac cat gcc ctc ata agg agc ctt gca aaa tca cat gaa gtg cta	1920

Leu Asp His Ala Leu Ile Arg Ser Leu Ala Lys Ser His Glu Val Leu  
 625 630 635 640  
 atc act gtc gaa gaa gga tca att gga ggt ttt gga tct cat gtt gtt 1968  
 Ile Thr Val Glu Glu Gly Ser Ile Gly Gly Phe Gly Ser His Val Val  
 645 650 655  
 cag ttc atg gcc tta gat ggg ctt ctt gat ggc aag ttg aag tgg aga 2016  
 Gln Phe Met Ala Leu Asp Gly Leu Leu Asp Gly Lys Leu Lys Trp Arg  
 660 665 670  
 cca ata gtt ctt cct gat cga tac att gac cat gga tct cct gtt gat 2064  
 Pro Ile Val Leu Pro Asp Arg Tyr Ile Asp His Gly Ser Pro Val Asp  
 675 680 685  
 cag ttg gcg gaa gct ggc cta aca cca tct cac att gca gca aca gta 2112  
 Gln Leu Ala Glu Ala Gly Leu Thr Pro Ser His Ile Ala Ala Thr Val  
 690 695 700  
 ttt aac ata ctt gga caa acc aga gag gct cta gag gtc atg aca taa 2160  
 Phe Asn Ile Leu Gly Gln Thr Arg Glu Ala Leu Glu Val Met Thr  
 705 710 715

&lt;210&gt; 58

&lt;211&gt; 719

&lt;212&gt; PRT

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 58

Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly  
 1 5 10 15  
 Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile  
 20 25 30  
 His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu  
 35 40 45  
 Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly  
 50 55 60  
 Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn  
 65 70 75 80

Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu  
85 90 95

Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly  
100 105 110

Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu  
115 120 125

His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly  
130 135 140

His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met  
145 150 155 160

Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser  
165 170 175

Glu Ser Glu Tyr Asp Cys Phe Gly Thr Gly His Ser Ser Thr Thr Ile  
180 185 190

Ser Ala Gly Leu Gly Met Ala Val Gly Arg Asp Leu Lys Gly Arg Asn  
195 200 205

Asn Asn Val Ile Ala Val Ile Gly Asp Gly Ala Met Thr Ala Gly Gln  
210 215 220

Ala Tyr Glu Ala Met Asn Asn Ala Gly Tyr Leu Asp Ser Asp Met Ile  
225 230 235 240

Val Ile Leu Asn Asp Asn Arg Gln Val Ser Leu Pro Thr Ala Thr Leu  
245 250 255

Asp Gly Pro Val Ala Pro Val Gly Ala Leu Ser Ser Ala Leu Ser Arg  
260 265 270

Leu Gln Ser Asn Arg Pro Leu Arg Glu Leu Arg Glu Val Ala Lys Gly  
275 280 285

Val Thr Lys Gln Ile Gly Gly Pro Met His Glu Leu Ala Ala Lys Val  
290 295 300

Asp Glu Tyr Ala Arg Gly Met Ile Ser Gly Ser Gly Ser Thr Leu Phe  
305 310 315 320

Glu Glu Leu Gly Leu Tyr Tyr Ile Gly Pro Val Asp Gly His Asn Ile  
325 330 335

Asp Asp Leu Ile Ala Ile Leu Lys Glu Val Arg Ser Thr Lys Thr Thr  
340 345 350

Gly Pro Val Leu Ile His Val Val Thr Glu Lys Gly Arg Gly Tyr Pro  
355 360 365

Tyr Ala Glu Arg Ala Ala Asp Lys Tyr His Gly Val Ala Lys Phe Asp  
370 375 380

Pro Ala Thr Gly Lys Gln Phe Lys Ala Ser Ala Lys Thr Gln Ser Tyr  
385 390 395 400

Thr Thr Tyr Phe Ala Glu Ala Leu Ile Ala Glu Ala Glu Ala Asp Lys  
405 410 415

Asp Ile Val Ala Ile His Ala Ala Met Gly Gly Gly Thr Gly Met Asn  
420 425 430

Leu Phe His Arg Arg Phe Pro Thr Arg Cys Phe Asp Val Gly Ile Ala  
435 440 445

Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Cys Glu Gly Ile  
450 455 460

Lys Pro Phe Cys Ala Ile Tyr Ser Ser Phe Met Gln Arg Ala Tyr Asp  
465 470 475 480

Gln Val Val His Asp Val Asp Leu Gln Lys Leu Pro Val Arg Phe Ala  
485 490 495

Met Asp Arg Ala Gly Leu Val Gly Ala Asp Gly Pro Thr His Cys Gly  
500 505 510

Ala Phe Asp Val Thr Tyr Met Ala Cys Leu Pro Asn Met Val Val Met  
515 520 525

Ala Pro Ser Asp Glu Ala Glu Leu Phe His Met Val Ala Thr Ala Ala  
 530 535 540

Ala Ile Asp Asp Arg Pro Ser Cys Phe Arg Tyr Pro Arg Gly Asn Gly  
 545 550 555 560

Ile Gly Val Glu Leu Pro Ala Gly Asn Lys Gly Ile Pro Leu Glu Val  
 565 570 575

Gly Lys Gly Arg Ile Leu Ile Glu Gly Glu Arg Val Ala Leu Leu Gly  
 580 585 590

Tyr Gly Ser Ala Val Gln Asn Cys Leu Asp Ala Ala Ile Val Leu Glu  
 595 600 605

Ser Arg Gly Leu Gln Val Thr Val Ala Asp Ala Arg Phe Cys Lys Pro  
 610 615 620

Leu Asp His Ala Leu Ile Arg Ser Leu Ala Lys Ser His Glu Val Leu  
 625 630 635 640

Ile Thr Val Glu Glu Gly Ser Ile Gly Gly Phe Gly Ser His Val Val  
 645 650 655

Gln Phe Met Ala Leu Asp Gly Leu Leu Asp Gly Lys Leu Lys Trp Arg  
 660 665 670

Pro Ile Val Leu Pro Asp Arg Tyr Ile Asp His Gly Ser Pro Val Asp  
 675 680 685

Gln Leu Ala Glu Ala Gly Leu Thr Pro Ser His Ile Ala Ala Thr Val  
 690 695 700

Phe Asn Ile Leu Gly Gln Thr Arg Glu Ala Leu Glu Val Met Thr  
 705 710 715

<210> 59

<211> 1434

<212> DNA

<213> Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1434)

&lt;223&gt;

&lt;400&gt; 59

atg atg aca tta aac tca cta tct cca gct gaa tcc aaa gct att tct	48
Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser	
1 5 10 15	
ttc ttg gat acc tcc agg ttc aat cca atc cct aaa ctc tca ggt ggg	96
Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly	
20 25 30	
ttt agt ttg agg agg agg aat caa ggg aga ggt ttt gga aaa ggt gtt	144
Phe Ser Leu Arg Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val	
35 40 45	
aag tgt tca gtg aaa gtg cag cag caa caa cct cct cca gca tgg	192
Lys Cys Ser Val Lys Val Gln Gln Gln Gln Pro Pro Pro Ala Trp	
50 55 60	
cct ggg aga gct gtc cct gag gcg cct cgt caa tct tgg gat gga cca	240
Pro Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro	
65 70 75 80	
aaa ccc atc tct atc gtt gga tct act ggt tct att ggc act cag aca	288
Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr	
85 90 95	
ttg gat att gtg gct gag aat cct gac aaa ttc aga gtt gtg gct cta	336
Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu	
100 105 110	
gct gct ggt tcg aat gtt act cta ctt gct gat cag gta agg aga ttt	384
Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe	
115 120 125	
aag cct gca ttg gtt gct gtt aga aac gag tca ctg att aat gag ctt	432
Lys Pro Ala Leu Val Ala Val Arg Asn Glu Ser Leu Ile Asn Glu Leu	
130 135 140	
aaa gag gct tta gct gat ttg gac tat aaa ctc gag att att cca gga	480
Lys Glu Ala Leu Ala Asp Leu Asp Tyr Lys Leu Glu Ile Ile Pro Gly	
145 150 155 160	
gag caa gga gtg att gag gtt gcc cga cat cct gaa gct gta acc gtt	528
Glu Gln Gly Val Ile Glu Val Ala Arg His Pro Glu Ala Val Thr Val	
165 170 175	



gtt acc gga ata gta ggt tgt gcg gga cta aag cct acg gtt gct gca	576
Val Thr Gly Ile Val Gly Cys Ala Gly Leu Lys Pro Thr Val Ala Ala	
180 185 190	
att gaa gca gga aag gac att gct ctt gca aac aaa gag aca tta atc	624
Ile Glu Ala Gly Lys Asp Ile Ala Leu Ala Asn Lys Glu Thr Leu Ile	
195 200 205	
gca ggt ggt cct ttc gtg ctt ccg ctt gcc aac aaa cat aat gta aag	672
Ala Gly Gly Pro Phe Val Leu Pro Leu Ala Asn Lys His Asn Val Lys	
210 215 220	
att ctt ccg gca gat tca gaa cat tct gcc ata ttt cag tgt att caa	720
Ile Leu Pro Ala Asp Ser Glu His Ser Ala Ile Phe Gln Cys Ile Gln	
225 230 235 240	
ggt ttg cct gaa ggc gct ctg cgc aag ata atc ttg act gca tct ggt	768
Gly Leu Pro Glu Gly Ala Leu Arg Lys Ile Ile Leu Thr Ala Ser Gly	
245 250 255	
gga gct ttt agg gat tgg cct gtc gaa aag cta aag gaa gtt aaa gta	816
Gly Ala Phe Arg Asp Trp Pro Val Glu Lys Leu Lys Glu Val Lys Val	
260 265 270	
gcg gat gcg ttg aag cat cca aac tgg aac atg gga aag aaa atc act	864
Ala Asp Ala Leu Lys His Pro Asn Trp Asn Met Gly Lys Lys Ile Thr	
275 280 285	
gtg gac tct gct acg ctt ttc aac aag ggt ctt gag gtc att gaa gcg	912
Val Asp Ser Ala Thr Leu Phe Asn Lys Gly Leu Glu Val Ile Glu Ala	
290 295 300	
cat tat ttg ttt gga gct gag tat gac gat ata gag att gtc att cat	960
His Tyr Leu Phe Gly Ala Glu Tyr Asp Asp Ile Glu Ile Val Ile His	
305 310 315 320	
ccg caa agt atc ata cat tcc atg att gaa aca cag gat tca tct gtg	1008
Pro Gln Ser Ile Ile His Ser Met Ile Glu Thr Gln Asp Ser Ser Val	
325 330 335	
ctt gct caa ttg ggt tgg cct gat atg cgt tta ccg att ctc tac acc	1056
Leu Ala Gln Leu Gly Trp Pro Asp Met Arg Leu Pro Ile Leu Tyr Thr	
340 345 350	
atg tca tgg ccc gat aga gtt cct tgt tct gaa gta act tgg cca aga	1104
Met Ser Trp Pro Asp Arg Val Pro Cys Ser Glu Val Thr Trp Pro Arg	
355 360 365	
ctt gac ctt tgc aaa ctc ggt tca ttg act ttc aag aaa cca gac aat	1152
Leu Asp Leu Cys Lys Leu Gly Ser Leu Thr Phe Lys Lys Pro Asp Asn	
370 375 380	
gtg aaa tac cca tcc atg gat ctt gct tat gct gct gga cga gct gga	1200
Val Lys Tyr Pro Ser Met Asp Leu Ala Tyr Ala Ala Gly Arg Ala Gly	
385 390 395 400	

ggc aca atg act gga gtt ctc agc gcc gcc aat gag aaa gct gtt gaa 1248  
 Gly Thr Met Thr Gly Val Leu Ser Ala Ala Asn Glu Lys Ala Val Glu  
 405 410 415

atg ttc att gat gaa aag ata agc tat ttg gat atc ttc aag gtt gtg 1296  
 Met Phe Ile Asp Glu Lys Ile Ser Tyr Leu Asp Ile Phe Lys Val Val  
 420 425 430

gaa tta aca tgc gat aaa cat cga aac gag ttg gta aca tca ccg tct 1344  
 Glu Leu Thr Cys Asp Lys His Arg Asn Glu Leu Val Thr Ser Pro Ser  
 435 440 445

ctt gaa gag att gtt cac tat gac ttg tgg gca cgt gaa tat gcc gcg 1392  
 Leu Glu Glu Ile Val His Tyr Asp Leu Trp Ala Arg Glu Tyr Ala Ala  
 450 455 460

aat gtg cag ctt tct tct ggt gct agg cca gtt cat gca tga 1434  
 Asn Val Gln Leu Ser Ser Gly Ala Arg Pro Val His Ala  
 465 470 475

<210> 60

<211> 477

<212> PRT

<213> Arabidopsis thaliana

<400> 60 [ ]

Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser  
 1 5 10 15

Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly  
 20 25 30

Phe Ser Leu Arg Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val  
 35 40 45

Lys Cys Ser Val Lys Val Gln Gln Gln Gln Gln Pro Pro Pro Ala Trp  
 50 55 60

Pro Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro  
 65 70 75 80

Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr  
 85 90 95

Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu  
100 105 110

Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe  
115 120 125

Lys Pro Ala Leu Val Ala Val Arg Asn Glu Ser Leu Ile Asn Glu Leu  
130 135 140

Lys Glu Ala Leu Ala Asp Leu Asp Tyr Lys Leu Glu Ile Ile Pro Gly  
145 150 155 160

Glu Gln Gly Val Ile Glu Val Ala Arg His Pro Glu Ala Val Thr Val  
165 170 175

Val Thr Gly Ile Val Gly Cys Ala Gly Leu Lys Pro Thr Val Ala Ala  
180 185 190

Ile Glu Ala Gly Lys Asp Ile Ala Leu Ala Asn Lys Glu Thr Leu Ile  
195 200 205

Ala Gly Gly Pro Phe Val Leu Pro Leu Ala Asn Lys His Asn Val Lys  
210 215 220

Ile Leu Pro Ala Asp Ser Glu His Ser Ala Ile Phe Gln Cys Ile Gln  
225 230 235 240

Gly Leu Pro Glu Gly Ala Leu Arg Lys Ile Ile Leu Thr Ala Ser Gly  
245 250 255

Gly Ala Phe Arg Asp Trp Pro Val Glu Lys Leu Lys Glu Val Lys Val  
260 265 270

Ala Asp Ala Leu Lys His Pro Asn Trp Asn Met Gly Lys Lys Ile Thr  
275 280 285

Val Asp Ser Ala Thr Leu Phe Asn Lys Gly Leu Glu Val Ile Glu Ala  
290 295 300

His Tyr Leu Phe Gly Ala Glu Tyr Asp Asp Ile Glu Ile Val Ile His  
305 310 315 320

Pro Gln Ser Ile Ile His Ser Met Ile Glu Thr Gln Asp Ser Ser Val  
325 330 335

Leu Ala Gln Leu Gly Trp Pro Asp Met Arg Leu Pro Ile Leu Tyr Thr  
340 345 350

Met Ser Trp Pro Asp Arg Val Pro Cys Ser Glu Val Thr Trp Pro Arg  
355 360 365

Leu Asp Leu Cys Lys Leu Gly Ser Leu Thr Phe Lys Lys Pro Asp Asn  
370 375 380

Val Lys Tyr Pro Ser Met Asp Leu Ala Tyr Ala Ala Gly Arg Ala Gly  
385 390 395 400

Gly Thr Met Thr Gly Val Leu Ser Ala Ala Asn Glu Lys Ala Val Glu  
405 410 415

Met Phe Ile Asp Glu Lys Ile Ser Tyr Leu Asp Ile Phe Lys Val Val  
420 425 430

Glu Leu Thr Cys Asp Lys His Arg Asn Glu Leu Val Thr Ser Pro Ser  
435 440 445

Leu Glu Glu Ile Val His Tyr Asp Leu Trp Ala Arg Glu Tyr Ala Ala  
450 455 460

Asn Val Gln Leu Ser Ser Gly Ala Arg Pro Val His Ala  
465 470 475

<210> 61

<211> 884

<212> DNA

<213> Adonis palaestina clone ApIPI28

<220>

<221> CDS

&lt;222&gt; (180)..(884)

&lt;223&gt;

&lt;400&gt; 61

```

cgtcgatcag gattaatcct ttatatagta tcttctccac caccactaaa acattatcag      60
cttcgtgttc ttctcccgct gttcatcttc agcagcggtg tcgtactctt tctattttctt    120
cttccatcac taacagtcct cgccgagggt tgaatcggct gttcgcctca acgtcgact      179
atg ggt gaa gtc gct gat gct ggt atg gat gcc gtc cag aag cgg ctt      227
Met Gly Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu
1          5          10          15
atg ttc gac gat gaa tgt att ttg gtg gat gag aat gac aag gtc gtc      275
Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val
          20          25          30
gga cat gat tcc aaa tac aac tgt cat ttg atg gaa aag ata gag gca      323
Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala
          35          40          45
gaa aac ttg ctt cac aga gcc ttc agt gtt ttc tta ttc aac tca aaa      371
Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
          50          55          60
tac gag ttg ctt ctt cag caa cga tct gca acg aag gta aca ttc ccg      419
Tyr Glu Leu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro
65          70          75          80
ctc gta tgg aca aac acc tgt tgc agc cat ccc ctc ttc cgt gat tcc      467
Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser
          85          90          95
gaa ctc ata gaa gaa aat ttt ctc ggg gta cga aac gct gca caa agg      515
Glu Leu Ile Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg
          100          105          110
aag ctt tta gac gag cta ggc att cca gct gaa gac gta cca gtt gat      563
Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp
          115          120          125
gaa ttc act cct ctt ggt cgc att ctt tac aaa gct cca tct gac gga      611
Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly
          130          135          140
aaa tgg gga gag cac gaa ctg gac tat ctt ctg ttt att gtc cga gat      659
Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp
145          150          155          160
gtg aaa tac gat cca aac cca gat gaa gtt gct gac gct aag tac gtt      707
Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val
          165          170          175

```

aat cgc gag gag ttg aaa gag ata ctg aga aaa gct gat gca ggt gaa 755  
 Asn Arg Glu Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu  
 180 185 190

gag gga ata aag ttg tct cct tgg ttt aga ttg gtt gtg gat aac ttt 803  
 Glu Gly Ile Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe  
 195 200 205

ttg ttc aag tgg tgg gat cat gta gag gag ggg aag att aag gac gtc 851  
 Leu Phe Lys Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val  
 210 215 220

gcc gac atg aaa act atc cac aag ttg act taa 884  
 Ala Asp Met Lys Thr Ile His Lys Leu Thr.  
 225 230

<210> 62

<211> 234

<212> PRT

<213> Adonis palaestina clone ApIPI28

<400> 62

Met Gly Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu  
 1 5 10 15

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val  
 20 25 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala  
 35 40 45

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys  
 50 55 60

Tyr Glu Leu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro  
 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser  
 85 90 95

Glu Leu Ile Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg  
 100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp  
 115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly  
 130 135 140

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp  
 145 150 155 160

Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val  
 165 170 175

Asn Arg Glu Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu  
 180 185 190

Glu Gly Ile Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe  
 195 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val  
 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr  
 225 230

<210> 63

<211> 1402

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (52)..(1317)

<223>

<400> 63

aagtctttgct ctcttttggtt tacttttcctc tgttttcgat ccatttagaa a atg tta

Met Leu

1

ttc acg agg agt gtt gct cgg att tct tct aag ttt ctg aga aac cgt	105
Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg Asn Arg	
5 10 15	
agc ttc tat ggc tcc tct caa tct ctc gcc tct cat cgg ttc gca atc	153
Ser Phe Tyr Gly Ser Ser Gln Ser Leu Ala Ser His Arg Phe Ala Ile	
20 25 30	
att ccc gat cag ggt cac tct tgt tct gac tct cca cac aag ggt tac	201
Ile Pro Asp Gln Gly His Ser Cys Ser Asp Ser Pro His Lys Gly Tyr	
35 40 45 50	
gtt tgc aga aca act tat tca ttg aaa tct ccg gtt ttt ggt gga ttt	249
Val Cys Arg Thr Thr Tyr Ser Leu Lys Ser Pro Val Phe Gly Gly Phe	
55 60 65	
agt cat caa ctc tat cac cag agt agc tcc ttg gtt gag gag gag ctt	297
Ser His Gln Leu Tyr His Gln Ser Ser Ser Leu Val Glu Glu Glu Leu	
70 75 80	
gac cca ttt tcg ctt gtt gcc gat gag ctg tca ctt ctt agt aat aag	345
Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Leu Leu Ser Asn Lys	
85 90 95	
ttg aga gag atg gta ctt gcc gag gtt cca aag ctt gcc tct gct gct	393
Leu Arg Glu Met Val Leu Ala Glu Val Pro Lys Leu Ala Ser Ala Ala	
100 105 110	
gag tac ttc ttc aaa agg ggt gtg caa gga aaa cag ttt cgt tca act	441
Glu Tyr Phe Phe Lys Arg Gly Val Gln Gly Lys Gln Phe Arg Ser Thr	
115 120 125 130	
att ttg ctg ctg atg gcg aca gct ctg gat gta cga gtt cca gaa gca	489
Ile Leu Leu Leu Met Ala Thr Ala Leu Asp Val Arg Val Pro Glu Ala	
135 140 145	
ttg att ggg gaa tca aca gat ata gtc aca tca gaa tta cgc gta agg	537
Leu Ile Gly Glu Ser Thr Asp Ile Val Thr Ser Glu Leu Arg Val Arg	
150 155 160	
caa cgg ggt att gct gaa atc act gaa atg ata cac gtc gca agt cta	585
Gln Arg Gly Ile Ala Glu Ile Thr Glu Met Ile His Val Ala Ser Leu	
165 170 175	
ctg cac gat gat gtc ttg gat gat gcc gat aca agg cgt ggt gtt ggt	633
Leu His Asp Asp Val Leu Asp Asp Ala Asp Thr Arg Arg Gly Val Gly	
180 185 190	
tcc tta aat gtt gta atg ggt aac aag atg tcg gta tta gca gga gac	681
Ser Leu Asn Val Val Met Gly Asn Lys Met Ser Val Leu Ala Gly Asp	
195 200 205 210	
ttc ttg ctc tcc cgg gct tgt ggg gct ctc gct gct tta aag aac aca	729





&lt;210&gt; 64

&lt;211&gt; 422

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 64

Met Leu Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg  
 1 5 10 15

Asn Arg Ser Phe Tyr Gly Ser Ser Gln Ser Leu Ala Ser His Arg Phe  
 20 25 30

Ala Ile Ile Pro Asp Gln Gly His Ser Cys Ser Asp Ser Pro His Lys  
 35 40 45

Gly Tyr Val Cys Arg Thr Thr Tyr Ser Leu Lys Ser Pro Val Phe Gly  
 50 55 60

Gly Phe Ser His Gln Leu Tyr His Gln Ser Ser Ser Leu Val Glu Glu  
 65 70 75 80

Glu Leu Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Leu Leu Ser  
 85 90 95

Asn Lys Leu Arg Glu Met Val Leu Ala Glu Val Pro Lys Leu Ala Ser  
 100 105 110

Ala Ala Glu Tyr Phe Phe Lys Arg Gly Val Gln Gly Lys Gln Phe Arg  
 115 120 125

Ser Thr Ile Leu Leu Leu Met Ala Thr Ala Leu Asp Val Arg Val Pro  
 130 135 140

Glu Ala Leu Ile Gly Glu Ser Thr Asp Ile Val Thr Ser Glu Leu Arg  
 145 150 155 160

Val Arg Gln Arg Gly Ile Ala Glu Ile Thr Glu Met Ile His Val Ala  
 165 170 175

Ser Leu Leu His Asp Asp Val Leu Asp Asp Ala Asp Thr Arg Arg Gly  
 180 185 190

Val Gly Ser Leu Asn Val Val Met Gly Asn Lys Met Ser Val Leu Ala  
 195 200 205

Gly Asp Phe Leu Leu Ser Arg Ala Cys Gly Ala Leu Ala Ala Leu Lys  
 210 215 220

Asn Thr Glu Val Val Ala Leu Leu Ala Thr Ala Val Glu His Leu Val  
 225 230 235 240

Thr Gly Glu Thr Met Glu Ile Thr Ser Ser Thr Glu Gln Arg Tyr Ser  
 245 250 255

Met Asp Tyr Tyr Met Gln Lys Thr Tyr Tyr Lys Thr Ala Ser Leu Ile  
 260 265 270

Ser Asn Ser Cys Lys Ala Val Ala Val Leu Thr Gly Gln Thr Ala Glu  
 275 280 285

Val Ala Val Leu Ala Phe Glu Tyr Gly Arg Asn Leu Gly Leu Ala Phe  
 290 295 300

Gln Leu Ile Asp Asp Ile Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu  
 305 310 315 320

Gly Lys Gly Ser Leu Ser Asp Ile Arg His Gly Val Ile Thr Ala Pro  
 325 330 335

Ile Leu Phe Ala Met Glu Glu Phe Pro Gln Leu Arg Glu Val Val Asp  
 340 345 350

Gln Val Glu Lys Asp Pro Arg Asn Val Asp Ile Ala Leu Glu Tyr Leu  
 355 360 365

Gly Lys Ser Lys Gly Ile Gln Arg Ala Arg Glu Leu Ala Met Glu His  
 370 375 380

Ala Asn Leu Ala Ala Ala Ala Ile Gly Ser Leu Pro Glu Thr Asp Asn  
 385 390 395 400

Glu Asp Val Lys Arg Ser Arg Arg Ala Leu Ile Asp Leu Thr His Arg  
 405 410 415

Val Ile Thr Arg Asn Lys  
 420

<210> 65

<211> 1155

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1155)

<223>

<400> 65

atg agt gtg agt tgt tgt tgt agg aat ctg ggc aag aca ata aaa aag 48  
 Met Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys  
 1 5 10 15

gca ata cct tca cat cat ttg cat ctg aga agt ctt ggt ggg agt ctc 96  
 Ala Ile Pro Ser His His Leu His Leu Arg Ser Leu Gly Gly Ser Leu  
 20 25 30

tat cgt cgt cgt atc caa agc tct tca atg gag acc gat ctc aag tca 144  
 Tyr Arg Arg Arg Ile Gln Ser Ser Ser Met Glu Thr Asp Leu Lys Ser  
 35 40 45

acc ttt ctc aac gtt tat tct gtt ctc aag tct gac ctt ctt cat gac 192  
 Thr Phe Leu Asn Val Tyr Ser Val Leu Lys Ser Asp Leu Leu His Asp  
 50 55 60

cct tcc ttc gaa ttc acc aat gaa tct cgt ctc tgg gtt gat cgg atg 240  
 Pro Ser Phe Glu Phe Thr Asn Glu Ser Arg Leu Trp Val Asp Arg Met  
 65 70 75 80

ctg gac tac aat gta cgt gga ggg aaa ctc aat cgg ggt ctc tct gtt 288  
 Leu Asp Tyr Asn Val Arg Gly Gly Lys Leu Asn Arg Gly Leu Ser Val  
 85 90 95

gtt gac agt ttc aaa ctt ttg aag caa ggc aat gat ttg act gag caa 336

Val Asp Ser Phe Lys Leu Leu Lys Gln Gly Asn Asp Leu Thr Glu Gln	
100 105 110	
gag gtt ttc ctc tct tgt gct ctc ggt tgg tgc att gaa tgg ctc caa	384
Glu Val Phe Leu Ser Cys Ala Leu Gly Trp Cys Ile Glu Trp Leu Gln	
115 120 125	
gct tat ttc ctt gtg ctt gat gat att atg gat aac tct gtc act cgc	432
Ala Tyr Phe Leu Val Leu Asp Asp Ile Met Asp Asn Ser Val Thr Arg	
130 135 140	
cgt ggt caa cct tgc tgg ttc aga gtt cct cag gtt ggt atg gtt gcc	480
Arg Gly Gln Pro Cys Trp Phe Arg Val Pro Gln Val Gly Met Val Ala	
145 150 155 160	
atc aat gat ggg att cta ctt cgc aat cac atc cac agg att ctc aaa	528
Ile Asn Asp Gly Ile Leu Leu Arg Asn His Ile His Arg Ile Leu Lys	
165 170 175	
aag cat ttc cgt gat aag cct tac tat gtt gac ctt gtt gat ttg ttt	576
Lys His Phe Arg Asp Lys Pro Tyr Tyr Val Asp Leu Val Asp Leu Phe	
180 185 190	
aat gag gtt gag ttg caa aca gct tgt ggc cag atg ata gat ttg atc	624
Asn Glu Val Glu Leu Gln Thr Ala Cys Gly Gln Met Ile Asp Leu Ile	
195 200 205	
acc acc ttt gaa gga gaa aag gat ttg gcc aag tac tca ttg tca atc	672
Thr Thr Phe Glu Gly Glu Lys Asp Leu Ala Lys Tyr Ser Leu Ser Ile	
210 215 220	
cac cgt cgt att gtc cag tac aaa acg gct tat tac tca ttt tat ctc	720
His Arg Arg Ile Val Gln Tyr Lys Thr Ala Tyr Tyr Ser Phe Tyr Leu	
225 230 235 240	
cct gtt gct tgt gcg ttg ctt atg gcg ggc gaa aat ttg gaa aac cat	768
Pro Val Ala Cys Ala Leu Leu Met Ala Gly Glu Asn Leu Glu Asn His	
245 250 255	
att gac gtg aaa aat gtt ctt gtt gac atg gga atc tac ttc caa gtg	816
Ile Asp Val Lys Asn Val Leu Val Asp Met Gly Ile Tyr Phe Gln Val	
260 265 270	
cag gat gat tat ctg gat tgt ttt gct gat ccc gag acg ctt ggc aag	864
Gln Asp Asp Tyr Leu Asp Cys Phe Ala Asp Pro Glu Thr Leu Gly Lys	
275 280 285	
ata gga aca gat ata gaa gat ttc aaa tgc tcg tgg ttg gtg gtt aag	912
Ile Gly Thr Asp Ile Glu Asp Phe Lys Cys Ser Trp Leu Val Val Lys	
290 295 300	
gca tta gag cgc tgc agc gaa gaa caa act aag ata tta tat gag aac	960
Ala Leu Glu Arg Cys Ser Glu Glu Gln Thr Lys Ile Leu Tyr Glu Asn	
305 310 315 320	
tat ggt aaa ccc gac cca tcg aac gtt gct aaa gtg aag gat ctc tac	1008

Tyr Gly Lys Pro Asp Pro Ser Asn Val Ala Lys Val Lys Asp Leu Tyr  
 325 330 335  
 aaa gag ctg gat ctt gag gga gtt ttc atg gag tat gag agc aaa agc 1056  
 Lys Glu Leu Asp Leu Glu Gly Val Phe Met Glu Tyr Glu Ser Lys Ser  
 340 345 350  
 tac gag aag ctg act gga gcg att gag gga cac caa agt aaa gca atc 1104  
 Tyr Glu Lys Leu Thr Gly Ala Ile Glu Gly His Gln Ser Lys Ala Ile  
 355 360 365  
 caa gca gtg cta aaa tcc ttc ttg gct aag atc tac aag agg cag aag 1152  
 Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys  
 370 375 380  
 tag 1155

&lt;210&gt; 66

&lt;211&gt; 384

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 66

Met Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys  
 1 5 10 15  
 Ala Ile Pro Ser His His Leu His Leu Arg Ser Leu Gly Gly Ser Leu  
 20 25 30  
 Tyr Arg Arg Arg Ile Gln Ser Ser Ser Met Glu Thr Asp Leu Lys Ser  
 35 40 45  
 Thr Phe Leu Asn Val Tyr Ser Val Leu Lys Ser Asp Leu Leu His Asp  
 50 55 60  
 Pro Ser Phe Glu Phe Thr Asn Glu Ser Arg Leu Trp Val Asp Arg Met  
 65 70 75 80  
 Leu Asp Tyr Asn Val Arg Gly Gly Lys Leu Asn Arg Gly Leu Ser Val  
 85 90 95  
 Val Asp Ser Phe Lys Leu Leu Lys Gln Gly Asn Asp Leu Thr Glu Gln  
 100 105 110

Glu Val Phe Leu Ser Cys Ala Leu Gly Trp Cys Ile Glu Trp Leu Gln  
115 120 125

Ala Tyr Phe Leu Val Leu Asp Asp Ile Met Asp Asn Ser Val Thr Arg  
130 135 140

Arg Gly Gln Pro Cys Trp Phe Arg Val Pro Gln Val Gly Met Val Ala  
145 150 155 160

Ile Asn Asp Gly Ile Leu Leu Arg Asn His Ile His Arg Ile Leu Lys  
165 170 175

Lys His Phe Arg Asp Lys Pro Tyr Tyr Val Asp Leu Val Asp Leu Phe  
180 185 190

Asn Glu Val Glu Leu Gln Thr Ala Cys Gly Gln Met Ile Asp Leu Ile  
195 200 205

Thr Thr Phe Glu Gly Glu Lys Asp Leu Ala Lys Tyr Ser Leu Ser Ile  
210 215 220

His Arg Arg Ile Val Gln Tyr Lys Thr Ala Tyr Tyr Ser Phe Tyr Leu  
225 230 235 240

Pro Val Ala Cys Ala Leu Leu Met Ala Gly Glu Asn Leu Glu Asn His  
245 250 255

Ile Asp Val Lys Asn Val Leu Val Asp Met Gly Ile Tyr Phe Gln Val  
260 265 270

Gln Asp Asp Tyr Leu Asp Cys Phe Ala Asp Pro Glu Thr Leu Gly Lys  
275 280 285

Ile Gly Thr Asp Ile Glu Asp Phe Lys Cys Ser Trp Leu Val Val Lys  
290 295 300

Ala Leu Glu Arg Cys Ser Glu Glu Gln Thr Lys Ile Leu Tyr Glu Asn  
305 310 315 320

Tyr Gly Lys Pro Asp Pro Ser Asn Val Ala Lys Val Lys Asp Leu Tyr  
325 330 335

Lys Glu Leu Asp Leu Glu Gly Val Phe Met Glu Tyr Glu Ser Lys Ser  
 340 345 350

Tyr Glu Lys Leu Thr Gly Ala Ile Glu Gly His Gln Ser Lys Ala Ile  
 355 360 365

Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys  
 370 375 380

<210> 67

<211> 1101

<212> DNA

<213> Sinabs alba

<220>

<221> CDS

<222> (1)..(1101)

<223>

<400> 67

atg gct tct tca gtg act cct cta ggt tca tgg gtt ctt ctt cac cat 48  
 Met Ala Ser Ser Val Thr Pro Leu Gly Ser Trp Val Leu Leu His His  
 1 5 10 15

cat cct tca act atc tta acc caa tcc aga tcc aga tct cct cct tct 96  
 His Pro Ser Thr Ile Leu Thr Gln Ser Arg Ser Arg Ser Pro Pro Ser  
 20 25 30

ctc atc acc ctt aaa ccc atc tcc ctc act cca aaa cgc acc gtt tcg 144  
 Leu Ile Thr Leu Lys Pro Ile Ser Leu Thr Pro Lys Arg Thr Val Ser  
 35 40 45

tct tct tcc tcc tct tcc ctc atc acc aaa gaa gac aac aac ctc aaa 192  
 Ser Ser Ser Ser Ser Ser Leu Ile Thr Lys Glu Asp Asn Asn Leu Lys  
 50 55 60

tcc tct tcc tct tcc ttc gat ttc atg tct tac atc atc cgc aaa gcc 240  
 Ser Ser Ser Ser Ser Phe Asp Phe Met Ser Tyr Ile Ile Arg Lys Ala  
 65 70 75 80

gac tcc gtc aac aaa gcc tta gac tcc gcc gtc cct ctc cgg gag cca 288



Asp	Ser	Val	Asn	Lys	Ala	Leu	Asp	Ser	Ala	Val	Pro	Leu	Arg	Glu	Pro	
			85						90					95		
ctc	aag	atc	cac	gaa	gcg	atg	cgt	tac	tct	ctc	ctc	gcc	gga	gga	aaa	336
Leu	Lys	Ile	His	Glu	Ala	Met	Arg	Tyr	Ser	Leu	Leu	Ala	Gly	Gly	Lys	
			100					105					110			
cgc	gtc	aga	cca	gtt	ctc	tgc	atc	gcc	gcg	tgc	gag	cta	gtc	gga	gga	384
Arg	Val	Arg	Pro	Val	Leu	Cys	Ile	Ala	Ala	Cys	Glu	Leu	Val	Gly	Gly	
			115				120					125				
gaa	gag	tct	tta	gct	atg	ccg	gcg	cgt	tgc	gcc	gtg	gaa	atg	atc	cac	432
Glu	Glu	Ser	Leu	Ala	Met	Pro	Ala	Arg	Cys	Ala	Val	Glu	Met	Ile	His	
	130						135				140					
acc	atg	tcg	ttg	atc	cac	gac	gac	ttg	cct	tgt	atg	gat	aac	gac	gat	480
Thr	Met	Ser	Leu	Ile	His	Asp	Asp	Leu	Pro	Cys	Met	Asp	Asn	Asp	Asp	
145					150					155				160		
ctc	cgc	cgc	gga	aag	ccc	acg	aat	cac	aaa	gtt	tac	ggc	gaa	gac	gtg	528
Leu	Arg	Arg	Gly	Lys	Pro	Thr	Asn	His	Lys	Val	Tyr	Gly	Glu	Asp	Val	
			165						170					175		
gcg	gtt	tta	gcc	gga	gac	gcg	ctt	ctt	tcg	ttc	gcc	ttc	gag	cat	tta	576
Ala	Val	Leu	Ala	Gly	Asp	Ala	Leu	Leu	Ser	Phe	Ala	Phe	Glu	His	Leu	
			180					185					190			
gcg	tcg	gct	acg	agc	tcg	gag	gtt	tct	ccg	gcg	aga	gtg	gtt	aga	gct	624
Ala	Ser	Ala	Thr	Ser	Ser	Glu	Val	Ser	Pro	Ala	Arg	Val	Val	Arg	Ala	
		195					200					205				
gtg	gga	gag	ttg	gct	aaa	gcc	atc	ggc	acc	gaa	ggg	ctc	gtg	gcg	gga	672
Val	Gly	Glu	Leu	Ala	Lys	Ala	Ile	Gly	Thr	Glu	Gly	Leu	Val	Ala	Gly	
	210					215					220					
caa	gtg	gtg	gat	ata	agc	agt	gaa	ggg	ttg	gac	tta	aac	aac	gtc	gga	720
Gln	Val	Val	Asp	Ile	Ser	Ser	Glu	Gly	Leu	Asp	Leu	Asn	Asn	Val	Gly	
225					230				235					240		
ttg	gag	cat	ttg	aag	ttt	ata	cat	ttg	cat	aaa	acg	gcg	gcg	ttg	ctt	768
Leu	Glu	His	Leu	Lys	Phe	Ile	His	Leu	His	Lys	Thr	Ala	Ala	Leu	Leu	
			245					250						255		
gaa	gct	tca	gcg	gtt	ttg	ggt	ggg	atc	atc	ggt	gga	ggg	agt	gat	gaa	816
Glu	Ala	Ser	Ala	Val	Leu	Gly	Gly	Ile	Ile	Gly	Gly	Gly	Ser	Asp	Glu	
			260				265						270			
gag	atc	gag	agg	ctg	agg	aag	ttc	gcg	agg	tgt	att	ggg	ttg	ttg	ttt	864
Glu	Ile	Glu	Arg	Leu	Arg	Lys	Phe	Ala	Arg	Cys	Ile	Gly	Leu	Leu	Phe	
		275					280					285				
cag	gtg	gtt	gat	gat	atc	ttg	gac	gtg	acg	aaa	tcg	tct	caa	gaa	ctg	912
Gln	Val	Val	Asp	Asp	Ile	Leu	Asp	Val	Thr	Lys	Ser	Ser	Gln	Glu	Leu	
	290					295					300					
ggg	aaa	acc	gct	ggg	aaa	gat	ttg	att	gct	gat	aag	ttg	act	tat	ccg	960

Gly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp Lys Leu Thr Tyr Pro  
 305 310 315 320

aag ctc atg ggt ttg gag aaa tcg aga gag ttc gct gag aag ttg aat 1008  
 Lys Leu Met Gly Leu Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn  
 325 330 335

aca gag gca cgt gat cag ctt tta ggg ttt gat tcc gac aag gtt gct 1056  
 Thr Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp Ser Asp Lys Val Ala  
 340 345 350

cct ttg ttg gct ttg gct aat tac att gcc aat aga cag aac tga 1101  
 Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Asn Arg Gln Asn  
 355 360 365

&lt;210&gt; 68

&lt;211&gt; 366

&lt;212&gt; PRT

&lt;213&gt; Sinabs alba

&lt;400&gt; 68

Met Ala Ser Ser Val Thr Pro Leu Gly Ser Trp Val Leu Leu His His  
 1 5 10 15

His Pro Ser Thr Ile Leu Thr Gln Ser Arg Ser Arg Ser Pro Pro Ser  
 20 25 30

Leu Ile Thr Leu Lys Pro Ile Ser Leu Thr Pro Lys Arg Thr Val Ser  
 35 40 45

Ser Ser Ser Ser Ser Ser Leu Ile Thr Lys Glu Asp Asn Asn Leu Lys  
 50 55 60

Ser Ser Ser Ser Ser Phe Asp Phe Met Ser Tyr Ile Ile Arg Lys Ala  
 65 70 75 80

Asp Ser Val Asn Lys Ala Leu Asp Ser Ala Val Pro Leu Arg Glu Pro  
 85 90 95

Leu Lys Ile His Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys  
 100 105 110

Arg Val Arg Pro Val Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly  
 115 120 125

Glu Glu Ser Leu Ala Met Pro Ala Arg Cys Ala Val Glu Met Ile His  
 130 135 140

Thr Met Ser Leu Ile His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp  
 145 150 155 160

Leu Arg Arg Gly Lys Pro Thr Asn His Lys Val Tyr Gly Glu Asp Val  
 165 170 175

Ala Val Leu Ala Gly Asp Ala Leu Leu Ser Phe Ala Phe Glu His Leu  
 180 185 190

Ala Ser Ala Thr Ser Ser Glu Val Ser Pro Ala Arg Val Val Arg Ala  
 195 200 205

Val Gly Glu Leu Ala Lys Ala Ile Gly Thr Glu Gly Leu Val Ala Gly  
 210 215 220

Gln Val Val Asp Ile Ser Ser Glu Gly Leu Asp Leu Asn Asn Val Gly  
 225 230 235 240

Leu Glu His Leu Lys Phe Ile His Leu His Lys Thr Ala Ala Leu Leu  
 245 250 255

Glu Ala Ser Ala Val Leu Gly Gly Ile Ile Gly Gly Gly Ser Asp Glu  
 260 265 270

Glu Ile Glu Arg Leu Arg Lys Phe Ala Arg Cys Ile Gly Leu Leu Phe  
 275 280 285

Gln Val Val Asp Asp Ile Leu Asp Val Thr Lys Ser Ser Gln Glu Leu  
 290 295 300

Gly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp Lys Leu Thr Tyr Pro  
 305 310 315 320

Lys Leu Met Gly Leu Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn  
 325 330 335

Thr Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp Ser Asp Lys Val Ala  
 340 345 350

Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Asn Arg Gln Asn  
 355 360 365

<210> 69

<211> 930

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(930)

<223>

<400> 69

atg aat aat ccg tcg tta ctc aat cat gcg gtc gaa acg atg gca gtt 48  
 Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val  
 1 5 10 15

ggc tcg aaa agt ttt gcg aca gcc tca aag tta ttt gat gca aaa acc 96  
 Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr  
 20 25 30

cgg cgc agc gta ctg atg ctc tac gcc tgg tgc cgc cat tgt gac gat 144  
 Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp  
 35 40 45

gtt att gac gat cag acg ctg ggc ttt cag gcc cgg cag cct gcc tta 192  
 Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu  
 50 55 60

caa acg ccc gaa caa cgt ctg atg caa ctt gag atg aaa acg cgc cag 240  
 Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln  
 65 70 75 80

gcc tat gca gga tcg cag atg cac gaa ccg gcg ttt gcg gct ttt cag 288  
 Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln  
 85 90 95

gaa gtg gct atg gct cat gat atc gcc ccg gct tac gcg ttt gat cat 336  
 Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His  
 100 105 110

ctg gaa ggc ttc gcc atg gat gta cgc gaa gcg caa tac agc caa ctg	384
Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu	
115 120 125	
gat gat acg ctg cgc tat tgc tat cac gtt gca ggc gtt gtc ggc ttg	432
Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu	
130 135 140	
atg atg gcg caa atc atg ggc gtg cgg gat aac gcc acg ctg gac cgc	480
Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg	
145 150 155 160	
gcc tgt gac ctt ggg ctg gca ttt cag ttg acc aat att gct cgc gat	528
Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp	
165 170 175	
att gtg gac gat gcg cat gcg ggc cgc tgt tat ctg ccg gca agc tgg	576
Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp	
180 185 190	
ctg gag cat gaa ggt ctg aac aaa gag aat tat gcg gca cct gaa aac	624
Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn	
195 200 205	
cgt cag gcg ctg agc cgt atc gcc cgt cgt ttg gtg cag gaa gca gaa	672
Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val Gln Glu Ala Glu	
210 215 220	
cct tac tat ttg tct gcc aca gcc ggc ctg gca ggg ttg ccc ctg cgt	720
Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu Pro Leu Arg	
225 230 235 240	
tcc gcc tgg gca atc gct acg gcg aag cag gtt tac cgg aaa ata ggt	768
Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg Lys Ile Gly	
245 250 255	
gtc aaa gtt gaa cag gcc ggt cag caa gcc tgg gat cag cgg cag tca	816
Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln Arg Gln Ser	
260 265 270	
acg acc acg ccc gaa aaa tta acg ctg ctg ctg gcc gcc tct ggt cag	864
Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Leu Ala Ala Ser Gly Gln	
275 280 285	
gcc ctt act tcc cgg atg cgg gct cat cct ccc cgc cct gcg cat ctc	912
Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro Ala His Leu	
290 295 300	
tgg cag cgc ccg ctc tag	930
Trp Gln Arg Pro Leu	
305	

&lt;211&gt; 309

&lt;212&gt; PRT

&lt;213&gt; Erwinia uredovora

&lt;400&gt; 70

Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val  
 1 5 10 15

Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr  
 20 25 30

Arg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp  
 35 40 45

Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu  
 50 55 60

Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln  
 65 70 75 80

Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln  
 85 90 95

Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His  
 100 105 110

Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu  
 115 120 125

Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu  
 130 135 140

Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg  
 145 150 155 160

Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp  
 165 170 175

Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp  
 180 185 190

Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn  
195 200 205

Arg Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Val Gln Glu Ala Glu  
210 215 220

Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu Pro Leu Arg  
225 230 235 240

Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg Lys Ile Gly  
245 250 255

Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln Arg Gln Ser  
260 265 270

Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Leu Ala Ala Ser Gly Gln  
275 280 285

Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro Ala His Leu  
290 295 300

Trp Gln Arg Pro Leu  
305

<210> 71

<211> 1479

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(1479)

<223>

<400> 71

atg aaa cca act acg gta att ggt gca ggc ttc ggt ggc ctg gca ctg

48

Met	Lys	Pro	Thr	Thr	Val	Ile	Gly	Ala	Gly	Phe	Gly	Gly	Leu	Ala	Leu	
1				5					10					15		
gca	att	cgt	cta	caa	gct	gcg	ggg	atc	ccc	gtc	tta	ctg	ctt	gaa	caa	96
Ala	Ile	Arg	Leu	Gln	Ala	Ala	Gly	Ile	Pro	Val	Leu	Leu	Leu	Glu	Gln	
			20				25					30				
cgt	gat	aaa	ccc	ggc	ggg	cgg	gct	tat	gtc	tac	gag	gat	cag	ggg	ttt	144
Arg	Asp	Lys	Pro	Gly	Gly	Arg	Ala	Tyr	Val	Tyr	Glu	Asp	Gln	Gly	Phe	
		35				40					45					
acc	ttt	gat	gca	ggc	ccg	acg	gtt	atc	acc	gat	ccc	agt	gcc	att	gaa	192
Thr	Phe	Asp	Ala	Gly	Pro	Thr	Val	Ile	Thr	Asp	Pro	Ser	Ala	Ile	Glu	
	50				55					60						
gaa	ctg	ttt	gca	ctg	gca	gga	aaa	cag	tta	aaa	gag	tat	gtc	gaa	ctg	240
Glu	Leu	Phe	Ala	Leu	Ala	Gly	Lys	Gln	Leu	Lys	Glu	Tyr	Val	Glu	Leu	
65				70					75				80			
ctg	ccg	gtt	acg	ccg	ttt	tac	cgc	ctg	tgt	tgg	gag	tca	ggg	aag	gtc	288
Leu	Pro	Val	Thr	Pro	Phe	Tyr	Arg	Leu	Cys	Trp	Glu	Ser	Gly	Lys	Val	
			85					90					95			
ttt	aat	tac	gat	aac	gat	caa	acc	cgg	ctc	gaa	gcg	cag	att	cag	cag	336
Phe	Asn	Tyr	Asp	Asn	Asp	Gln	Thr	Arg	Leu	Glu	Ala	Gln	Ile	Gln	Gln	
			100					105					110			
ttt	aat	ccc	cgc	gat	gtc	gaa	ggg	tat	cgt	cag	ttt	ctg	gac	tat	tca	384
Phe	Asn	Pro	Arg	Asp	Val	Glu	Gly	Tyr	Arg	Gln	Phe	Leu	Asp	Tyr	Ser	
		115				120						125				
cgc	gcg	gtg	ttt	aaa	gaa	ggc	tat	cta	aag	ctc	ggg	act	gtc	cct	ttt	432
Arg	Ala	Val	Phe	Lys	Glu	Gly	Tyr	Leu	Lys	Leu	Gly	Thr	Val	Pro	Phe	
	130					135					140					
tta	tcg	ttc	aga	gac	atg	ctt	cgc	gcc	gca	cct	caa	ctg	gcg	aaa	ctg	480
Leu	Ser	Phe	Arg	Asp	Met	Leu	Arg	Ala	Ala	Pro	Gln	Leu	Ala	Lys	Leu	
145					150					155				160		
cag	gca	tgg	aga	agc	gtt	tac	agt	aag	gtt	gcc	agt	tac	atc	gaa	gat	528
Gln	Ala	Trp	Arg	Ser	Val	Tyr	Ser	Lys	Val	Ala	Ser	Tyr	Ile	Glu	Asp	
			165					170					175			
gaa	cat	ctg	cgc	cag	gcg	ttt	tct	ttc	cac	tcg	ctg	ttg	gtg	ggc	ggc	576
Glu	His	Leu	Arg	Gln	Ala	Phe	Ser	Phe	His	Ser	Leu	Leu	Val	Gly	Gly	
			180					185					190			
aat	ccc	ttc	gcc	acc	tca	tcc	att	tat	acg	ttg	ata	cac	gcg	ctg	gag	624
Asn	Pro	Phe	Ala	Thr	Ser	Ser	Ile	Tyr	Thr	Leu	Ile	His	Ala	Leu	Glu	
		195					200					205				
cgt	gag	tgg	ggc	gtc	tgg	ttt	ccg	cgt	ggc	ggc	acc	ggc	gca	tta	gtt	672
Arg	Glu	Trp	Gly	Val	Trp	Phe	Pro	Arg	Gly	Gly	Thr	Gly	Ala	Leu	Val	
	210					215					220					
cag	ggg	atg	ata	aag	ctg	ttt	cag	gat	ctg	ggg	gaa	gtc	gtg	tta		720



Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu 225 230 235 240	
aac gcc aga gtc agc cat atg gaa acg aca gga aac aag att gaa gcc Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala 245 250 255	768
gtg cat tta gag gac ggt cgc agg ttc ctg acg caa gcc gtc gcg tca Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser 260 265 270	816
aat gca gat gtg gtt cat acc tat cgc gac ctg tta agc cag cac cct Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro 275 280 285	864
gcc gcg gtt aag cag tcc aac aaa ctg cag act aag cgc atg agt aac Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn 290 295 300	912
tct ctg ttt gtg ctc tat ttt ggt ttg aat cac cat cat gat cag ctc Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu 305 310 315 320	960
gcg cat cac acg gtt tgt ttc ggc ccg cgt tac cgc gag ctg att gac Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp 325 330 335	1008
gaa att ttt aat cat gat ggc ctc gca gag gac ttc tca ctt tat ctg Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu 340 345 350	1056
cac gcg ccc tgt gtc acg gat tcg tca ctg gcg cct gaa ggt tgc ggc His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly 355 360 365	1104
agt tac tat gtg ttg gcg ccg gtg ccg cat tta ggc acc gcg aac ctc Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu 370 375 380	1152
gac tgg acg gtt gag ggg cca aaa cta cgc gac cgt att ttt gcg tac Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr 385 390 395 400	1200
ctt gag cag cat tac atg cct ggc tta cgg agt cag ctg gtc acg cac Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His 405 410 415	1248
cgg atg ttt acg ccg ttt gat ttt cgc gac cag ctt aat gcc tat cat Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His 420 425 430	1296
ggc tca gcc ttt tct gtg gag ccc gtt ctt acc cag agc gcc tgg ttt Gly Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe 435 440 445	1344
cgg ccg cat aac cgc gat aaa acc att act aat ctc tac ctg gtc ggc	1392

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly  
 450 455 460  
 gca ggc acg cat ccc ggc gca ggc att cct ggc gtc atc ggc tcg gca 1440  
 Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala  
 465 470 475 480  
 aaa gcg aca gca ggt ttg atg ctg gag gat ctg ata tga 1479  
 Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile  
 485 490

<210> 72

<211> 492

<212> PRT

<213> Erwinia uredovora

<400> 72

Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu  
 1 5 10 15

Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Leu Glu Gln  
 20 25 30

Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Glu Asp Gln Gly Phe  
 35 40 45

Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu  
 50 55 60

Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Glu Tyr Val Glu Leu  
 65 70 75 80

Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val  
 85 90 95

Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu Ala Gln Ile Gln Gln  
 100 105 110

Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln Phe Leu Asp Tyr Ser  
 115 120 125

Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe  
 130 135 140

Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu  
 145 150 155 160

Gln Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Ser Tyr Ile Glu Asp  
 165 170 175

Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly  
 180 185 190

Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu  
 195 200 205

Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val  
 210 215 220

Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Gly Glu Val Val Leu  
 225 230 235 240

Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala  
 245 250 255

Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser  
 260 265 270

Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro  
 275 280 285

Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn  
 290 295 300

Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu  
 305 310 315 320

Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp  
 325 330 335

Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu  
 340 345 350

His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly  
355 360 365

Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu  
370 375 380

Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr  
385 390 395 400

Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His  
405 410 415

Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His  
420 425 430

Gly Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe  
435 440 445

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly  
450 455 460

Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala  
465 470 475 480

Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile  
485 490

<210> 73

<211> 1725

<212> DNA

<213> Narcissus pseudonarcissus

<220>

<221> CDS

<222> (1) .. (1725)

<223>

&lt;400&gt; 73

atg gct tct tcc act tgt tta att cat tct tcc tct ttt ggg gtt gga	48
Met Ala Ser Ser Thr Cys Leu Ile His Ser Ser Ser Phe Gly Val Gly	
1 5 10 15	
 gga aag aaa gtg aag atg aac acg atg att cga tgc aag ttg ttt tca	96
Gly Lys Lys Val Lys Met Asn Thr Met Ile Arg Ser Lys Leu Phe Ser	
20 25 30	
 att cgg tgc gct ttg gac act aag gtg tct gat atg agc gtc aat gct	144
Ile Arg Ser Ala Leu Asp Thr Lys Val Ser Asp Met Ser Val Asn Ala	
35 40 45	
 cca aaa gga ttg ttt cca cca gag cct gag cac tac agg ggg cca aag	192
Pro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro Lys	
50 55 60	
 ctt aaa gtg gct atc att gga gct ggg ctc gct ggc atg tca act gca	240
Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr Ala	
65 70 75 80	
 gtg gag ctt ttg gat caa ggg cat gag gtt gac ata tat gaa tcc aga	288
Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser Arg	
85 90 95	
 caa ttt att ggt ggt aaa gtc ggt tct ttt gta gat aag cgt gga aac	336
Gln Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Lys Arg Gly Asn	
100 105 110	
 cat att gaa atg gga ctc cat gtg ttt ttt ggt tgc tat aac aat ctt	384
His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn Leu	
115 120 125	
 ttc aga ctt atg aaa aag gta ggt gca gat gaa aat tta ctg gtg aag	432
Phe Arg Leu Met Lys Lys Val Gly Ala Asp Glu Asn Leu Leu Val Lys	
130 135 140	
 gat cat act cat acc ttt gta aac cga ggt gga gaa att ggt gaa ctt	480
Asp His Thr His Thr Phe Val Asn Arg Gly Gly Glu Ile Gly Glu Leu	
145 150 155 160	
 gat ttc cga ctt ccg atg ggt gca cca tta cat ggt att cgt gca ttt	528
Asp Phe Arg Leu Pro Met Gly Ala Pro Leu His Gly Ile Arg Ala Phe	
165 170 175	
 cta aca act aat caa ctg aag cct tat gat aaa gca agg aat gct gtg	576
Leu Thr Thr Asn Gln Leu Lys Pro Tyr Asp Lys Ala Arg Asn Ala Val	
180 185 190	
 gct ctt gcc ctt agc cca gtt gta cgt gct ctt att gat cca aat ggt	624
Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Ile Asp Pro Asn Gly	
195 200 205	
 gca atg cag gat ata agg aac tta gat aat att agc ttt tct gat tgg	672
Ala Met Gln Asp Ile Arg Asn Leu Asp Asn Ile Ser Phe Ser Asp Trp	
210 215 220	

ttc tta tcc aaa ggc ggt acc cgc atg agc atc caa agg atg tgg gat	720
Phe Leu Ser Lys Gly Gly Thr Arg Met Ser Ile Gln Arg Met Trp Asp	
225 230 235 240	
cca gtt gct tat gcc ctc gga ttt att gac tgt gat aat atc agt gcc	768
Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser Ala	
245 250 255	
cgt tgt atg ctt act ata ttt tct cta ttt gct act aag aca gaa gct	816
Arg Cys Met Leu Thr Ile Phe Ser Leu Phe Ala Thr Lys Thr Glu Ala	
260 265 270	
tct ctg ttg cgt atg ttg aag ggt tcg cct gat gtt tac tta agc ggt	864
Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser Gly	
275 280 285	
cct ata aga aag tat att aca gat aaa ggt gga agg ttt cac cta agg	912
Pro Ile Arg Lys Tyr Ile Thr Asp Lys Gly Gly Arg Phe His Leu Arg	
290 295 300	
tgg ggg tgt aga gag ata ctt tat gat gaa cta tca aat ggc gac aca	960
Trp Gly Cys Arg Glu Ile Leu Tyr Asp Glu Leu Ser Asn Gly Asp Thr	
305 310 315 320	
tat atc aca ggc att gca atg tcg aag gct acc aat aaa aaa ctt gtg	1008
Tyr Ile Thr Gly Ile Ala Met Ser Lys Ala Thr Asn Lys Lys Leu Val	
325 330 335	
aaa gct gac gtg tat gtt gca gca tgt gat gtt cct gga ata aaa agg	1056
Lys Ala Asp Val Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys Arg	
340 345 350	
ttg atc cca tcg gag tgg aga gaa tgg gat cta ttt gac aat atc tat	1104
Leu Ile Pro Ser Glu Trp Arg Glu Trp Asp Leu Phe Asp Asn Ile Tyr	
355 360 365	
aaa cta gtt gga gtt cca gtt gtc act gtt cag ctt agg tac aat ggt	1152
Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn Gly	
370 375 380	
tgg gtg aca gag atg caa gat ctg gaa aaa tca agg cag ttg aga gct	1200
Trp Val Thr Glu Met Gln Asp Leu Glu Lys Ser Arg Gln Leu Arg Ala	
385 390 395 400	
gca gta gga ttg gat aat ctt ctt tat act cca gat gca gac ttt tct	1248
Ala Val Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe Ser	
405 410 415	
tgt ttt tct gat ctt gca ctc tcg tcg cct gaa gat tat tat att gaa	1296
Cys Phe Ser Asp Leu Ala Leu Ser Ser Pro Glu Asp Tyr Tyr Ile Glu	
420 425 430	
gga caa ggg tcc cta ata cag gct gtt ctc acg cca ggg gat cca tac	1344
Gly Gln Gly Ser Leu Ile Gln Ala Val Leu Thr Pro Gly Asp Pro Tyr	
435 440 445	

atg ccc cta cct aat gat gca att ata gaa aga gtt cgg aaa cag gtt 1392  
 Met Pro Leu Pro Asn Asp Ala Ile Ile Glu Arg Val Arg Lys Gln Val  
 450 455 460

ttg gat tta ttc cca tcc tct caa ggc ctg gaa gtt cta tgg tct tcg 1440  
 Leu Asp Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Leu Trp Ser Ser  
 465 470 475 480

gtg gtt aaa atc gga caa tcc cta tat cgg gag ggg cct gga aag gac 1488  
 Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys Asp  
 485 490 495

cca ttc aga cct gat cag aag aca cca gta aaa aat ttc ttc ctt gca 1536  
 Pro Phe Arg Pro Asp Gln Lys Thr Pro Val Lys Asn Phe Phe Leu Ala  
 500 505 510

ggt tca tac acc aaa cag gat tac att gac agt atg gaa gga gcg acc 1584  
 Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala Thr  
 515 520 525

cta tcg ggg aga caa gca gct gca tat atc tgc agc gcc ggt gaa gat 1632  
 Leu Ser Gly Arg Gln Ala Ala Tyr Ile Cys Ser Ala Gly Glu Asp  
 530 535 540

ctg gca gca ctt cgc aag aag atc gct gct gat cat cca gag caa ctg 1680  
 Leu Ala Ala Leu Arg Lys Lys Ile Ala Ala Asp His Pro Glu Gln Leu  
 545 550 555 560

atc aac aaa gat tct aac gtg tcg gat gaa ctg agt ctc gta taa 1725  
 Ile Asn Lys Asp Ser Asn Val Ser Asp Glu Leu Ser Leu Val  
 565 570

&lt;210&gt; 74

&lt;211&gt; 574

&lt;212&gt; PRT

&lt;213&gt; Narcissus pseudonarcissus

&lt;400&gt; 74

Met Ala Ser Ser Thr Cys Leu Ile His Ser Ser Ser Phe Gly Val Gly  
 1 5 10 15

Gly Lys Lys Val Lys Met Asn Thr Met Ile Arg Ser Lys Leu Phe Ser  
 20 25 30

Ile Arg Ser Ala Leu Asp Thr Lys Val Ser Asp Met Ser Val Asn Ala  
 35 40 45

Pro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro Lys  
50 55 60

Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr Ala  
65 70 75 80

Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser Arg  
85 90 95

Gln Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Lys Arg Gly Asn  
100 105 110

His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn Leu  
115 120 125

Phe Arg Leu Met Lys Lys Val Gly Ala Asp Glu Asn Leu Leu Val Lys  
130 135 140

Asp His Thr His Thr Phe Val Asn Arg Gly Gly Glu Ile Gly Glu Leu  
145 150 155 160

Asp Phe Arg Leu Pro Met Gly Ala Pro Leu His Gly Ile Arg Ala Phe  
165 170 175

Leu Thr Thr Asn Gln Leu Lys Pro Tyr Asp Lys Ala Arg Asn Ala Val  
180 185 190

Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Ile Asp Pro Asn Gly  
195 200 205

Ala Met Gln Asp Ile Arg Asn Leu Asp Asn Ile Ser Phe Ser Asp Trp  
210 215 220

Phe Leu Ser Lys Gly Gly Thr Arg Met Ser Ile Gln Arg Met Trp Asp  
225 230 235 240

Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser Ala  
245 250 255

Arg Cys Met Leu Thr Ile Phe Ser Leu Phe Ala Thr Lys Thr Glu Ala  
260 265 270



Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser Gly  
 275 280 285

Pro Ile Arg Lys Tyr Ile Thr Asp Lys Gly Gly Arg Phe His Leu Arg  
 290 295 300

Trp Gly Cys Arg Glu Ile Leu Tyr Asp Glu Leu Ser Asn Gly Asp Thr  
 305 310 315 320

Tyr Ile Thr Gly Ile Ala Met Ser Lys Ala Thr Asn Lys Lys Leu Val  
 325 330 335

Lys Ala Asp Val Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys Arg  
 340 345 350

Leu Ile Pro Ser Glu Trp Arg Glu Trp Asp Leu Phe Asp Asn Ile Tyr  
 355 360 365

Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn Gly  
 370 375 380

Trp Val Thr Glu Met Gln Asp Leu Glu Lys Ser Arg Gln Leu Arg Ala  
 385 390 395 400

Ala Val Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe Ser  
 405 410 415

Cys Phe Ser Asp Leu Ala Leu Ser Ser Pro Glu Asp Tyr Tyr Ile Glu  
 420 425 430

Gly Gln Gly Ser Leu Ile Gln Ala Val Leu Thr Pro Gly Asp Pro Tyr  
 435 440 445

Met Pro Leu Pro Asn Asp Ala Ile Ile Glu Arg Val Arg Lys Gln Val  
 450 455 460

Leu Asp Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Leu Trp Ser Ser  
 465 470 475 480

Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys Asp  
 485 490 495

Pro Phe Arg Pro Asp Gln Lys Thr Pro Val Lys Asn Phe Phe Leu Ala  
 500 505 510

Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala Thr  
 515 520 525

Leu Ser Gly Arg Gln Ala Ala Tyr Ile Cys Ser Ala Gly Glu Asp  
 530 535 540

Leu Ala Ala Leu Arg Lys Lys Ile Ala Ala Asp His Pro Glu Gln Leu  
 545 550 555 560

Ile Asn Lys Asp Ser Asn Val Ser Asp Glu Leu Ser Leu Val  
 565 570

<210> 75

<211> 1848

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(1848)

<223>

<400> 75

atg tgt acc ttg agt ttt atg tat cct aat tca ctt ctt gat ggt acc 48  
 Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr  
 1 5 10 15

tgc aag act gta gct ttg ggt gat agc aaa cca aga tac aat aaa cag 96  
 Cys Lys Thr Val Ala Leu Gly Asp Ser Lys Pro Arg Tyr Asn Lys Gln  
 20 25 30

aga agt tct tgt ttt gac cct ttg ata att gga aat tgt act gat cag 144  
 Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln  
 35 40 45

cag cag ctt tgt ggc ttg agt tgg ggg gtg gac aag gct aag gga aga 192

Gln	Gln	Leu	Cys	Gly	Leu	Ser	Trp	Gly	Val	Asp	Lys	Ala	Lys	Gly	Arg	
50						55					60					
aga	ggg	ggt	act	gtt	tcc	aat	ttg	aaa	gca	gtt	gta	gat	gta	gac	aaa	240
Arg	Gly	Gly	Thr	Val	Ser	Asn	Leu	Lys	Ala	Val	Val	Asp	Val	Asp	Lys	
65					70				75						80	
aga	gtg	gag	agc	tat	ggc	agt	agt	gat	gta	gaa	gga	aat	gag	agt	ggc	288
Arg	Val	Glu	Ser	Tyr	Gly	Ser	Ser	Asp	Val	Glu	Gly	Asn	Glu	Ser	Gly	
				85					90						95	
agc	tat	gat	gcc	att	gtt	ata	ggc	tca	gga	ata	ggc	gga	ttg	gtg	gca	336
Ser	Tyr	Asp	Ala	Ile	Val	Ile	Gly	Ser	Gly	Ile	Gly	Gly	Leu	Val	Ala	
			100					105					110			
gcg	acg	cag	ctg	gcg	gtt	aag	gga	gct	aag	gtt	tta	gtt	ctg	gag	aag	384
Ala	Thr	Gln	Leu	Ala	Val	Lys	Gly	Ala	Lys	Val	Leu	Val	Leu	Glu	Lys	
		115					120						125			
tat	gtt	att	cct	ggc	gga	agc	tct	ggc	ttt	tac	gag	agg	gat	ggc	tat	432
Tyr	Val	Ile	Pro	Gly	Gly	Ser	Ser	Gly	Phe	Tyr	Glu	Arg	Asp	Gly	Tyr	
	130					135					140					
aag	ttt	gat	gtt	ggc	tca	tca	gtg	atg	ttt	gga	ttc	agt	gat	aag	gga	480
Lys	Phe	Asp	Val	Gly	Ser	Ser	Val	Met	Phe	Gly	Phe	Ser	Asp	Lys	Gly	
145					150				155						160	
aac	ctc	aat	tta	att	act	caa	gca	ttg	gca	gca	gta	gga	cgt	aaa	tta	528
Asn	Leu	Asn	Leu	Ile	Thr	Gln	Ala	Leu	Ala	Ala	Val	Gly	Arg	Lys	Leu	
				165				170						175		
gaa	gtt	ata	cct	gac	cca	aca	act	gta	cat	ttc	cac	ctg	cca	aat	gac	576
Glu	Val	Ile	Pro	Asp	Pro	Thr	Thr	Val	His	Phe	His	Leu	Pro	Asn	Asp	
			180					185					190			
ctt	tct	gtt	cgt	ata	cac	cga	gag	tat	gat	gac	ttc	att	gaa	gag	ctt	624
Leu	Ser	Val	Arg	Ile	His	Arg	Glu	Tyr	Asp	Asp	Phe	Ile	Glu	Glu	Leu	
		195					200					205				
gtg	agt	aaa	ttt	cca	cat	gaa	aag	gaa	ggg	att	atc	aaa	ttt	tac	agt	672
Val	Ser	Lys	Phe	Pro	His	Glu	Lys	Glu	Gly	Ile	Ile	Lys	Phe	Tyr	Ser	
	210					215					220					
gaa	tgc	tgg	aag	atc	ttt	aat	tct	ctg	aat	tca	ttg	gaa	ctg	aag	tct	720
Glu	Cys	Trp	Lys	Ile	Phe	Asn	Ser	Leu	Asn	Ser	Leu	Glu	Leu	Lys	Ser	
225					230					235					240	
ttg	gag	gaa	ccc	atc	tac	ctt	ttt	ggc	cag	ttc	ttt	aag	aag	ccc	ctt	768
Leu	Glu	Glu	Pro	Ile	Tyr	Leu	Phe	Gly	Gln	Phe	Phe	Lys	Lys	Pro	Leu	
				245				250						255		
gaa	tgc	ttg	act	ctt	gcc	tac	tat	ttg	ccc	cag	aat	gct	ggc	agc	atc	816
Glu	Cys	Leu	Thr	Leu	Ala	Tyr	Tyr	Leu	Pro	Gln	Asn	Ala	Gly	Ser	Ile	
			260					265					270			
gct	cgg	aag	tat	ata	aga	gat	cct	ggg	ttg	ctg	tct	ttt	ata	gat	gca	864

Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala	
275 280 285	
gag tgc ttt atc gtg agt aca gtt aat gca tta caa aca cca atg atc	912
Glu Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gln Thr Pro Met Ile	
290 295 300	
aat gca agc atg gtt cta tgt gac aga cat ttt ggc gga atc aac tac	960
Asn Ala Ser Met Val Leu Cys Asp Arg His Phe Gly Gly Ile Asn Tyr	
305 310 315 320	
ccc gtt ggt gga gtt ggc gag atc gcc aaa tcc tta gca aaa ggc ttg	1008
Pro Val Gly Gly Val Gly Glu Ile Ala Lys Ser Leu Ala Lys Gly Leu	
325 330 335	
gat gat cac gga agt cag ata ctt tat agg gca aat gtt aca agt atc	1056
Asp Asp His Gly Ser Gln Ile Leu Tyr Arg Ala Asn Val Thr Ser Ile	
340 345 350	
att ttg gac aat ggc aaa gct gtg gga gtg aag ctt tct gac ggg agg	1104
Ile Leu Asp Asn Gly Lys Ala Val Gly Val Lys Leu Ser Asp Gly Arg	
355 360 365	
aag ttt tat gct aaa acc ata gta tcg aat gct acc aga tgg gat act	1152
Lys Phe Tyr Ala Lys Thr Ile Val Ser Asn Ala Thr Arg Trp Asp Thr	
370 375 380	
ttt gga aag ctt tta aaa gct gag aat ctg cca aaa gaa gaa gaa aat	1200
Phe Gly Lys Leu Leu Lys Ala Glu Asn Leu Pro Lys Glu Glu Glu Asn	
385 390 395 400	
ttc cag aaa gct tat gta aaa gca cct tct ttt ctt tct att cat atg	1248
Phe Gln Lys Ala Tyr Val Lys Ala Pro Ser Phe Leu Ser Ile His Met	
405 410 415	
gga gtt aaa gca gat gta ctc cca cca gac aca gat tgt cac cat ttt	1296
Gly Val Lys Ala Asp Val Leu Pro Pro Asp Thr Asp Cys His His Phe	
420 425 430	
gtc ctc gag gat gat tgg aca aat ttg gag aaa cca tat gga agt ata	1344
Val Leu Glu Asp Asp Trp Thr Asn Leu Glu Lys Pro Tyr Gly Ser Ile	
435 440 445	
ttc ttg agt att cca aca gtt ctt gat tcc tca ttg gcc cca gaa gga	1392
Phe Leu Ser Ile Pro Thr Val Leu Asp Ser Ser Leu Ala Pro Glu Gly	
450 455 460	
cac cat att ctt cac att ttt aca aca tcg agc att gaa gat tgg gag	1440
His His Ile Leu His Ile Phe Thr Thr Ser Ser Ile Glu Asp Trp Glu	
465 470 475 480	
gga ctc tct ccg aaa gac tat gaa gcg aag aaa gag gtt gtt gct gaa	1488
Gly Leu Ser Pro Lys Asp Tyr Glu Ala Lys Lys Glu Val Val Ala Glu	
485 490 495	
agg att ata agc aga ctt gaa aaa aca ctc ttc cca ggg ctt aag tca	1536

Arg Ile Ile Ser Arg Leu Glu Lys Thr Leu Phe Pro Gly Leu Lys Ser	
500	505 510
tct att ctc ttt aag gag gtg gga act cca aag acc cac aga cga tac	1584
Ser Ile Leu Phe Lys Glu Val Gly Thr Pro Lys Thr His Arg Arg Tyr	
515	520 525
ctt gct cgt gat agt ggt acc tat gga cca atg cca cgc gga aca cct	1632
Leu Ala Arg Asp Ser Gly Thr Tyr Gly Pro Met Pro Arg Gly Thr Pro	
530	535 540
aag gga ctc ctg gga atg cct ttc aat acc act gct ata gat ggt cta	1680
Lys Gly Leu Leu Gly Met Pro Phe Asn Thr Thr Ala Ile Asp Gly Leu	
545	550 555 560
tat tgt gtt ggc gat agt tgc ttc cca gga caa ggt gtt ata gct gta	1728
Tyr Cys Val Gly Asp Ser Cys Phe Pro Gly Gln Gly Val Ile Ala Val	
565	570 575
gcc ttt tca gga gta atg tgc gct cat cgt gtt gca gct gac tta ggg	1776
Ala Phe Ser Gly Val Met Cys Ala His Arg Val Ala Ala Asp Leu Gly	
580	585 590
ttt gaa aaa aaa tca gat gtg ctg gac agt gct ctt ctt aga cta ctt	1824
Phe Glu Lys Lys Ser Asp Val Leu Asp Ser Ala Leu Leu Arg Leu Leu	
595	600 605
ggt tgg tta agg aca cta gca tga	1848
Gly Trp Leu Arg Thr Leu Ala	
610	615

&lt;210&gt; 76

&lt;211&gt; 615

&lt;212&gt; PRT

&lt;213&gt; Lycopersicon esculentum

&lt;400&gt; 76

Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr
1 5 10 15

Cys Lys Thr Val Ala Leu Gly Asp Ser Lys Pro Arg Tyr Asn Lys Gln
20 25 30

Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln
35 40 45

Gln Gln Leu Cys Gly Leu Ser Trp Gly Val Asp Lys Ala Lys Gly Arg  
 50 55 60

Arg Gly Gly Thr Val Ser Asn Leu Lys Ala Val Val Asp Val Asp Lys  
 65 70 75 80

Arg Val Glu Ser Tyr Gly Ser Ser Asp Val Glu Gly Asn Glu Ser Gly  
 85 90 95

Ser Tyr Asp Ala Ile Val Ile Gly Ser Gly Ile Gly Gly Leu Val Ala  
 100 105 110

Ala Thr Gln Leu Ala Val Lys Gly Ala Lys Val Leu Val Leu Glu Lys  
 115 120 125

Tyr Val Ile Pro Gly Gly Ser Ser Gly Phe Tyr Glu Arg Asp Gly Tyr  
 130 135 140

Lys Phe Asp Val Gly Ser Ser Val Met Phe Gly Phe Ser Asp Lys Gly  
 145 150 155 160

Asn Leu Asn Leu Ile Thr Gln Ala Leu Ala Ala Val Gly Arg Lys Leu  
 165 170 175

Glu Val Ile Pro Asp Pro Thr Thr Val His Phe His Leu Pro Asn Asp  
 180 185 190

Leu Ser Val Arg Ile His Arg Glu Tyr Asp Asp Phe Ile Glu Glu Leu  
 195 200 205

Val Ser Lys Phe Pro His Glu Lys Glu Gly Ile Ile Lys Phe Tyr Ser  
 210 215 220

Glu Cys Trp Lys Ile Phe Asn Ser Leu Asn Ser Leu Glu Leu Lys Ser  
 225 230 235 240

Leu Glu Glu Pro Ile Tyr Leu Phe Gly Gln Phe Phe Lys Lys Pro Leu  
 245 250 255

Glu Cys Leu Thr Leu Ala Tyr Tyr Leu Pro Gln Asn Ala Gly Ser Ile  
 260 265 270

Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala  
275 280 285

Glu Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gln Thr Pro Met Ile  
290 295 300

Asn Ala Ser Met Val Leu Cys Asp Arg His Phe Gly Gly Ile Asn Tyr  
305 310 315 320

Pro Val Gly Gly Val Gly Glu Ile Ala Lys Ser Leu Ala Lys Gly Leu  
325 330 335

Asp Asp His Gly Ser Gln Ile Leu Tyr Arg Ala Asn Val Thr Ser Ile  
340 345 350

Ile Leu Asp Asn Gly Lys Ala Val Gly Val Lys Leu Ser Asp Gly Arg  
355 360 365

Lys Phe Tyr Ala Lys Thr Ile Val Ser Asn Ala Thr Arg Trp Asp Thr  
370 375 380

Phe Gly Lys Leu Leu Lys Ala Glu Asn Leu Pro Lys Glu Glu Glu Asn  
385 390 395 400

Phe Gln Lys Ala Tyr Val Lys Ala Pro Ser Phe Leu Ser Ile His Met  
405 410 415

Gly Val Lys Ala Asp Val Leu Pro Pro Asp Thr Asp Cys His His Phe  
420 425 430

Val Leu Glu Asp Asp Trp Thr Asn Leu Glu Lys Pro Tyr Gly Ser Ile  
435 440 445

Phe Leu Ser Ile Pro Thr Val Leu Asp Ser Ser Leu Ala Pro Glu Gly  
450 455 460

His His Ile Leu His Ile Phe Thr Thr Ser Ser Ile Glu Asp Trp Glu  
465 470 475 480

Gly Leu Ser Pro Lys Asp Tyr Glu Ala Lys Lys Glu Val Val Ala Glu  
485 490 495

Arg Ile Ile Ser Arg Leu Glu Lys Thr Leu Phe Pro Gly Leu Lys Ser  
 500 505 510

Ser Ile Leu Phe Lys Glu Val Gly Thr Pro Lys Thr His Arg Arg Tyr  
 515 520 525

Leu Ala Arg Asp Ser Gly Thr Tyr Gly Pro Met Pro Arg Gly Thr Pro  
 530 535 540

Lys Gly Leu Leu Gly Met Pro Phe Asn Thr Thr Ala Ile Asp Gly Leu  
 545 550 555 560

Tyr Cys Val Gly Asp Ser Cys Phe Pro Gly Gln Gly Val Ile Ala Val  
 565 570 575

Ala Phe Ser Gly Val Met Cys Ala His Arg Val Ala Ala Asp Leu Gly  
 580 585 590

Phe Glu Lys Lys Ser Asp Val Leu Asp Ser Ala Leu Leu Arg Leu Leu  
 595 600 605

Gly Trp Leu Arg Thr Leu Ala  
 610 615

<210> 77

<211> 1233

<212> DNA

<213> Tagetes erecta

<220>

<221> CDS

<222> (1)..(1233)

<223>

<400> 77

atg gcc aca cac aaa ctc ctt caa ttc acc acc aat ctc cca cca tct  
 Met Ala Thr His Lys Leu Leu Gln Phe Thr Thr Asn Leu Pro Pro Ser  
 1 5 10 15



tct tct tca atc tct act ggc tgt tca ctc tcc ccc ttc ttc ctc aaa	96
Ser Ser Ser Ile Ser Thr Gly Cys Ser Leu Ser Pro Phe Phe Leu Lys	
20 25 30	
tca tct tct cat tcc cct aac cct cgc cga cac cgc cgc tcc gcc gta	144
Ser Ser Ser His Ser Pro Asn Pro Arg Arg His Arg Arg Ser Ala Val	
35 40 45	
tgc tgc tct ttc gcc tca ctc gac tct gca aaa atc aaa gtc gtt ggc	192
Cys Cys Ser Phe Ala Ser Leu Asp Ser Ala Lys Ile Lys Val Val Gly	
50 55 60	
gtc ggt ggt ggt ggc aac aat gcc gtt aac cgc atg att ggt agc ggc	240
Val Gly Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Gly Ser Gly	
65 70 75 80	
tta cag ggt gtt gat ttt tac gcc att aac acg gac tca caa gcg ctt	288
Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu	
85 90 95	
ctg caa tct gtt gca cat aac cct att caa att ggg gag ctt ttg act	336
Leu Gln Ser Val Ala His Asn Pro Ile Gln Ile Gly Glu Leu Leu Thr	
100 105 110	
cgt gga tta ggt act ggt ggg aac ccg ctt ttg gga gaa cag gct gcg	384
Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala	
115 120 125	
gag gag tcg aag gaa gcg att ggg aat gcg ctt aaa ggg tcg gat ctt	432
Glu Glu Ser Lys Glu Ala Ile Gly Asn Ala Leu Lys Gly Ser Asp Leu	
130 135 140	
gtg ttt ata aca gca ggt atg ggt ggt ggg acg ggt tcg ggt gct gct	480
Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala	
145 150 155 160	
cca gtt gta gcg cag ata gcg aaa gaa gca ggg tat tta act gtt ggt	528
Pro Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly	
165 170 175	
gtt gta acg tac cca ttc agc ttt gaa ggc cgt aaa aga tca gta cag	576
Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln	
180 185 190	
gcg tta gag gct att gag aag ctg caa aag aac gtt gac aca ctt ata	624
Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile	
195 200 205	
gtg att cca aat gac cgt ttg ctg gat att gct gat gaa aac acg cct	672
Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Asn Thr Pro	
210 215 220	
ctt cag gat gct ttt ctt ctt gct gat gat gta ctc cgc caa gga gtt	720
Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val	
225 230 235 240	

caa gga atc tca gat ata att aca ata cct ggg ctg gta aat gtg gac	768
Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp	
245 250 255	
ttt gca gac gtt aaa gca gtc atg aaa gat tct gga act gca atg ctt	816
Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu	
260 265 270	
ggg gtc ggt gtt tcc tca agt aaa aac cga gct gaa gaa gca gct gaa	864
Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu	
275 280 285	
caa gca act ctt gct cct ttg att gga tca tca att caa tct gct aca	912
Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr	
290 295 300	
ggg gtt gtt tat aat att acc gga ggg aag gac ata act cta caa gaa	960
Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu	
305 310 315 320	
gtc aac agg gtt tct cag gtg gta aca agt ttg gca gat cca tca gca	1008
Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala	
325 330 335	
aac att ata ttc ggg gca gtg gta gat gag aga tac aac ggg gag att	1056
Asn Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile	
340 345 350	
cat gtg acc att gtt gct act ggc ttt gcc cag tcg ttt cag aaa tct	1104
His Val Thr Ile Val Ala Thr Gly Phe Ala Gln Ser Phe Gln Lys Ser	
355 360 365	
ctt ctt gct gac ccg aaa gga gca aaa ctt gtt gat aga aat caa gaa	1152
Leu Leu Ala Asp Pro Lys Gly Ala Lys Leu Val Asp Arg Asn Gln Glu	
370 375 380	
cct aca caa cct ttg act tcc gcg aga tct ttg aca aca cct tct cct	1200
Pro Thr Gln Pro Leu Thr Ser Ala Arg Ser Leu Thr Thr Pro Ser Pro	
385 390 395 400	
gct ccg tct cgg tct agg aaa ctc ttc ttt taa	1233
Ala Pro Ser Arg Ser Arg Lys Leu Phe Phe	
405 410	

&lt;210&gt; 78

&lt;211&gt; 410

&lt;212&gt; PRT

&lt;213&gt; Tagetes erecta

&lt;400&gt; 78

Met Ala Thr His Lys Leu Leu Gln Phe Thr Thr Asn Leu Pro Pro Ser  
1 5 10 15

Ser Ser Ser Ile Ser Thr Gly Cys Ser Leu Ser Pro Phe Phe Leu Lys  
20 25 30

Ser Ser Ser His Ser Pro Asn Pro Arg Arg His Arg Arg Ser Ala Val  
35 40 45

Cys Cys Ser Phe Ala Ser Leu Asp Ser Ala Lys Ile Lys Val Val Gly  
50 55 60

Val Gly Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Gly Ser Gly  
65 70 75 80

Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu  
85 90 95

Leu Gln Ser Val Ala His Asn Pro Ile Gln Ile Gly Glu Leu Leu Thr  
100 105 110

Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala  
115 120 125

Glu Glu Ser Lys Glu Ala Ile Gly Asn Ala Leu Lys Gly Ser Asp Leu  
130 135 140

Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala  
145 150 155 160

Pro Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly  
165 170 175

Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln  
180 185 190

Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile  
195 200 205

Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Asn Thr Pro  
210 215 220

Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val  
225 230 235 240

Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp  
245 250 255

Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu  
260 265 270

Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu  
275 280 285

Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr  
290 295 300

Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu  
305 310 315 320

Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala  
325 330 335

Asn Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile  
340 345 350

His Val Thr Ile Val Ala Thr Gly Phe Ala Gln Ser Phe Gln Lys Ser  
355 360 365

Leu Leu Ala Asp Pro Lys Gly Ala Lys Leu Val Asp Arg Asn Gln Glu  
370 375 380

Pro Thr Gln Pro Leu Thr Ser Ala Arg Ser Leu Thr Thr Pro Ser Pro  
385 390 395 400

Ala Pro Ser Arg Ser Arg Lys Leu Phe Phe  
405 410

<210> 79

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(33)

<223>

<400> 79

gcatgctcta gaccttataa agatattttg tga

33

<210> 80

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer\_bind

<222> (1)..(33)

<223>

<400> 80

gcatgcatct agaaatggtt cagtgtcaac cat

33

<210> 81

<211> 805

<212> DNA

<213> Nostoc sp. Strain PCC7120

<220>

<221> variation

<222> (1) .. (805)

<223>

<400> 81

```
gc atg cat ct ct ag aa at gg tt c ag t gt ca ac cat cat ct ct gc att c aga aa act gg t gt      60
tat t gt cat c gaca at caga gat gata aaaa at att aataa ggg tat attt att gc ct g ct      120
tt at ct tt att ttt at ggg ca att ag ttt aa tct t att act ct ca at ag at ac at cc ata a      180
tt ca ta ag ag ctt att ag gt at ag cc at gc ttt gg c ag ac ctt ctt ata t ac ag g ttt at      240
tt att act gc tcat gat gcc at gc ac gg cg tag ttt tat cc caaaa atccc aga ata aata      300
att ttt at ag g ta ag ct cact cta at ctt gt at gg act act cc tt tata aa gat tt att ga      360
aaaa ac att g gtt ac acc ac gg ac at cct g gt act gat tt ag acc ct gat tat ta ca at g      420
gt cat ccc ca aa act tct ttt ctt t gg tat c tac att ttt at ga ag tct tat tgg cg at gg a      480
cg caa at ttt cgg att ag t g at g at tttt c at gg act ta a aa at ct gg t g cat ata cc ag      540
aaa ata at ttt aatt ata ttt tgg at gata c ctt ct att ttt a ag tt c ag ta ca act at ttt      600
att ttt gg tac at tttt gc ct cataaaa agc tag a ag gt gg tt ata cta ac ccc att gt g      660
cg cg c ag tat ccc att acc t ctt tttt gg t ctt ttt gt tac tt gtt at c ac ttc gg ct acc      720
aca ag ga aca tc ac ga at ac cct ca act tc ctt gg t gg aa att acc t ga a gct ca caaaa      780
tat cttt ata ag gt ct ag ag cat gc      805
```

<210> 82

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (37)

<223>

<400> 82  
gcgcatgcat ctagaaatga tccagttaga acaacca

37

<210> 83

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(37)

<223>

<400> 83  
gcgcatgctc tagactatatt tgctttgtaa atttctg

37

<210> 84

<211> 792

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> misc\_feature

<222> (5)..(775)

<223>

<400> 84  
gcgcatgcat ctagaaatga tccagttaga acaaccactc agtcatcaag caaaactgac 60  
tccagtactg agaagtaaat ctcagtttaa ggggcttttc attgctattg tcattgttag 120  
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 180  
gatgttattg cctgttatac tatggcaaac atttttatat acgggattat ttattacatc 240

tcacgatgcc atgcatggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg 300  
aacattgacc ctatcccttt atgggtcttt accatatcaa aaactattga aaaaacattg 360  
gttacaccac cacaatccag caagctcaat agaccggat tttcacaatg gtaaacacca 420  
aagtttcttt gcttggtatt ttcattttat gaaaggttac tggagttggg ggcaaataat 480  
tgcgttgact attatttata actttgctaa atacatactc catatcccaa gtgataatct 540  
aacttacttt tgggtgctac cctcgctttt aagttcatta caattattct attttggtag 600  
ttttttaccc catagtgaac caataggggg ttatgttcag cctcattgtg cccaaacaat 660  
tagccgtcct atttggtggt catttatcac gtgctatcat tttggctacc acgaggaaca 720  
tcacgaatat cctcatattt cttggtggca gttaccagaa atttacaag caaaatagtc 780  
tagagcatgc gc 792

<210> 85

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (26)

<223>

<400> 85

gtcgaccctg ctttaatgag atatgc

26

<210> 86

<211> 27

<212> DNA

<213> Künstliche Sequenz



<220>

<221> Primer

<222> (1) .. (27)

<223>

<400> 86

ctcgagcttg gacaatcagt aaattga

27

<210> 87

<211> 210

<212> DNA

<213> Agrobacterium tumefaciens

<220>

<221> Terminator

<222> (1) .. (210)

<223>

<400> 87

gtcgaccctg ctttaatgag atatgcgaga cgcctatgat cgcgatgat ttgctttcaa 60

ttctgttgtag caggttgtaa aaaacctgag catgtgtagc tcagatcctt accgccggtt 120

tcggttcatt ctaatgaata tatcacccgt tactatcgta tttttatgaa taatattctc 180

cgttcaattt actgattgtc caagctcgag 210

<210> 88

<211> 1773

<212> DNA

<213> Petunia hybrida

<220>

<221> Promotor

<222> (1) .. (1773)

<223>

<400> 88

gagctctagc gcaatccttat gtggtacaaa tcttgattag tcgggaaaaa atgatgtggc	60
cctacaaatg gttggaggat gggagatttg gctctatcta gagttatgtg gttgttgaag	120
catttggtta ctctctgctg tggtagtgg catatccaca ttgtctcctt ccacttttat	180
gacaattacg tgaaagttaa gggttgtttt gtctattttt gtcgaggcct ttcttttcct	240
tccaggttgt tgaagatggc ccaattcgat tagaataatg ttttgagctt tagcatattc	300
tctctcgttt acacgattat agtaataatg atataggatg acagaagttg acacataaat	360
tttttattct ctccatttac tttaatccaa atctcaccta ccctaaactt ctttaatatg	420
tattcaatag tctatccgag taaattgtaa atttaacaac cattgataat attgacacct	480
actaacatat actagtaaag agaattataa catggcacat ataatttgat gcaaaatgag	540
tatgatgaaa tttaaaccca aaatctcttg attttgacag tgtcaccttg acttggtaac	600
taataagtca tgtttttagtg gcagaaagac aaactcatcc accaactgta tagcaataaa	660
aaatagaaga atcttcctga ggcaaagttt tggaaaaatt aagagtggct gagatttaat	720
ttcaacagga attagttcca cttaactttt aggttacgat acagtgtctaa ttaaataact	780
taattgtatt agatatctt tgcacctaaa aaatttataa actgaaaaaa ggtagcaatc	840
aaaataaaca aaaggacaaa ataagtgaag ggtacagcca ccaaccctgg cggctcactg	900
tttggttggtt aaaacgtaga cttacaccta ccaaatcta caactaaaat gaggcaataa	960
tactttgccc aaaattacca agaaaagaaa aagaaaggaa tcccttaata ttactctcct	1020
ccatttcaca ataaatatcc tagtttgact taaattagag tttaaaaaat gaaagacgac	1080
ttttaaaact tgtaatctaa aataaatcat agttaaatgt gtggctataa atcattgtat	1140
taacggtaaa gtggttaagt taaaagttaa ttgttttcaa atataaaatt gtactatcat	1200
tctttttgga atggactaat aagaaaacta tgacatccat tatggagcgg agggagtatc	1260
tccttttaac aataaccttt gtcccttcaa ttcaattatc agtatgcaaa cattaaaaat	1320
tattattgat gtttaagtacc acatcatcct taatgataga atcatcgtag aacgcttttc	1380
caggcacaca ttcaaactag ttagaccagt accacacatc gaatattcca gacttctttg	1440

tttgaatagt cgactacatt ggataatgga acttctcgaa ttaacttcga attagtcgag 1500  
cccaaaataa tatatacgtc ggggtggaaaa ctataaaatg ttgacaaaa atgtcaaatt 1560  
aatatatcaa tctgcaacaa ccttttcacc ttgagaacac agctgaaatt ttttaciaag 1620  
gtagttgggtg aagctagtca gcgaatccca ttaccttcca ctctaccta ccccttcac 1680  
caacaacaaa tttctgtaat ttaaaaacta gccaaaaaag aactctcttt taciaagagc 1740  
caaagactca atctttactt tcaagaaaag ctt 1773

<210> 89

<211> 29

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (29)

<223>

<400> 89

gagctctagc gcaatcttat gtggtacaa

29

<210> 90

<211> 29

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (29)

<223>

<400> 90  
aagcttttct tgaaagtaaa gattgagtc

29

<210> 91

<211> 39

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(39)

<223>

<400> 91  
gcgcatgcat ctagaaatga atttttgtga taaaccagt

39

<210> 92

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(37)

<223>

<400> 92  
gcgcatgctc tagattacga attgggttact gaattgt

37

<210> 93

<211> 819

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220>

<221> misc\_feature

<222> (5) .. (802)

<223>

<400> 93

gcgcatgcat ctagaaatga atttttgtga taaaccagtt agctattatg ttgcaataga	60
gcaattaagt gctaaagaag atactgtttg ggggctggtg attgtcatag taattattag	120
tctttgggta gctagtttgg cttttttact agctattaat tatgccaaag tcccaatttg	180
gttgatacct attgcaatag tttggcaaat gttcctttat acagggctat ttattactgc	240
acatgatgct atgcatgggt cagtttatcg taaaaatccc aaaattaata attttatcgg	300
ttcactagct gtacgcgttt acgctgtggt tccatatcaa cagatgttaa agaatcattg	360
cttacatcat cgtcacctg ctagcgaagt tgaccagat tttcatgatg gtaagagaac	420
aaacgctatt ttctggtatc tccatttcat gatagaatac tccagttggc aacagttaat	480
agtactaact atcctattta atttagctaa atacgttttg cacatccatc aaataaatct	540
catcttattt tggagtattc ctccaatttt aagtccatt caactgtttt atttcggaac	600
atttttgcct catcgagaac ccaagaaagg atatgtttat ccccatgca gccaaacaat	660
aaaattgcca acttttttgt catttatcgc ttgctaccac tttggttatc atgaagaaca	720
tcatgagtat ccccatgtac cttggtggca acttccatct gtatataagc agagagtatt	780
caacaattca gtaaccaatt cgtaatctag agcatgccc	819

<210> 94

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(33)

<223>

<400> 94  
gcgcatgcat ctagaaatgg cgatcgccat tat

33

<210> 95

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(32)

<223>

<400> 95  
gcgcatgctc tagatcacaa atttgattta ga

32

<210> 96

<211> 720

<212> DNA

<213> Nodularia spumigena NSOR10

<220>

<221> CDS

<222> (5)..(703)

&lt;223&gt;

&lt;400&gt; 96

```

gcgc atg cat cta gaa atg gcg atc gcc att att agt ata tgg gct atc      49
  Met His Leu Glu Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile
    1             5             10             15

agc cta ggt ttg tta ctt tat att gat ata tcc caa ttc aag ttt tgg      97
Ser Leu Gly Leu Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp
          20             25             30

atg ttg tta ccg ctc ata ttt tgg caa aca ttt tta tat acg gga tta      145
Met Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu
          35             40             45

ttt att aca gct cat gat gcc atg cat ggg gta gtt ttt ccc aaa aat      193
Phe Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn
          50             55             60

ccc aaa atc aac cat ttc att ggc tca ttg tgc ctg ttt ctt tat ggt      241
Pro Lys Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly
          65             70             75

ctt tta cct tat caa aaa ctt tta aaa aag cat tgg cta cat cac cat      289
Leu Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His His
          80             85             90             95

aat cca gcc agt gaa aca gat cca gat ttt cac aac ggg aag cag aaa      337
Asn Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys
          100             105             110

aac ttt ttt gct tgg tat tta tat ttt atg aag cgt tac tgg agt tgg      385
Asn Phe Phe Ala Trp Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp
          115             120             125

tta caa att atc aca tta atg att att tat aac tta cta aaa tat ata      433
Leu Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile
          130             135             140

tgg cat ttt cca gag gat aat atg act tat ttt tgg gta gtt ccc tca      481
Trp His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser
          145             150             155

att tta agt tct tta caa tta ttt tat ttt gga act ttt cta ccc cac      529
Ile Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His
          160             165             170             175

agt gag cct gta gaa ggt tat aaa gag cct cat cgt tcc caa act att      577
Ser Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile
          180             185             190

agc cgt ccc att tgg tgg tca ttt ata act tgt tac cat ttt ggt tat      625
Ser Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr
          195             200             205

```

cat tac gaa cat cat gaa tac ccc cat gtt cct tgg tgg caa tta cca 673  
 His Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro  
           210                                  215                                  220

gaa att tat aaa atg tct aaa tca aat ttg tgatctagag catgcgc 720  
 Glu Ile Tyr Lys Met Ser Lys Ser Asn Leu  
           225                                  230

<210> 97

<211> 233

<212> PRT

<213> Nodularia spumigena NSOR10

<400> 97

Met His Leu Glu Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser  
   1                                  5                                  10                                  15

Leu Gly Leu Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met  
                                   20                                  25                                  30

Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe  
                                   35                                  40                                  45

Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro  
                                   50                                  55                                  60

Lys Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu  
   65                                  70                                  75                                  80

Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His His Asn  
                                   85                                  90                                  95

Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn  
                                   100                                  105                                  110

Phe Phe Ala Trp Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu  
                                   115                                  120                                  125

Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp  
                                   130                                  135                                  140



His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile  
145 150 155 160

Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser  
165 170 175

Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser  
180 185 190

Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His  
195 200 205

Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu  
210 215 220

Ile Tyr Lys Met Ser Lys Ser Asn Leu  
225 230

<210> 98

<211> 37

<212> DNA

<213> Künstliche Sequenz

<220>

<221> primer

<222> (1)..(37)

<223>

<400> 98  
gagctcgata tctttgccag tattacaaca gcttata

37

<210> 99

<211> 31

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1)..(31)

<223>

<400> 99

cccgggttta ctgaaaaata acagtaaaac c

31

<210> 100

<211> 2096

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(2096)

<223>

<400> 100

gagctcgata tctttgccag tattacaaca gcttatatgt tgagcaggta aaagcttcaa 60

tgccctattc tttctacagt tatcaatgtt gtcgtcttaa tatctggtgt tcttctcgaa 120

atgtcaattg gcttgcagca cattgtcctc taatatccat tcaagcttct tagatgatga 180

aacatttgtc aaatttatta atttcatagt gttcagtctc aattctttag ctgggttcctc 240

atagtaaagt tgtctaatat gaaatgaaaa tgttctgtgt gttgtactaa taccttttca 300

tggttgctta tagaacgtcg atgaagagcc aaacagaaac tattttgggc tgcgatttct 360

gataccattg tatctgaatg ctgggtggga gctcatcaga agctttacaa tgggtcacat 420

atatggagcc ggtatgagga atgctgggaa tcagttgcgt ttcgcgtgct aggacttttc 480

cttcctggta tttctgccca cagcccagtt gattacgtga actccgtcag acttggaag 540

gagagaagta cccaaatgtc gtcttttttag aaatactttt gtcacaaaat agcgggggttt 600  
acagctacag aagatcatgc agaaggcgtc cagtttagtt tttgaaggtt gtttggagtt 660  
tatttatcta aagtaaactt aaatcagctt tttgtttatg agttcagtga actatatgtt 720  
caaataagac ttccctttgt agatatgtgt tttttttgtt gttgagcact ttgtgtgcat 780  
tgataaacc cccaacgtgt aatagctacc atacaagaga agtaactcgc actgtccatg 840  
tcttatgtgg ctcgactcag aaagcattca gggggattga taaccaccct ccaaaccaac 900  
tgaaccattg tgaataacca cccttcaaat caaccgagtc ctctgaagg acaaatatgt 960  
ggttttatat acattaaatt ttgtttttac atgcttcctc ttacttcttt agttttcttg 1020  
accatatctt gcgtttttcc cttctgtaat tgacactttt cttcaaacca tccagcaatg 1080  
tggaagcttg acgattttcc ttcagagtag aaattgaaaa gaatcaacta aaaaggatag 1140  
tccttcgatt tgatttccgg cttaaaaata aactaataag aatgagagag cgaataatag 1200  
aatattttga aattttaaag atattcaact atgttaaatt gcgttataaa tttcttaaatt 1260  
tagtagcacc taatagttaa gttctcaaaa gtcaaaacta ctacataatg tgctcatttt 1320  
tcacattaaa atgcctacat gatgtaaaag taaaactcgt agcattctac gtgttttact 1380  
caactcaaac atcctgttca ttttaataaa cgtacgatga gcttctctct ccaattttct 1440  
ttcttttttt ttttttaaaa aaatattttt ttttatatca atccaaatgg gctccaattt 1500  
atcataaatt aggtagaaac ttagatatta aagaaagaaa agggtttatc tcgcaagtgt 1560  
ggctatgggtg ggacgtgtca aattttggat tgtagccaaa catgagattt gatttaaagg 1620  
gaattggcca aatcaccgaa agcaggcatc ttcatacaa attagtttgt ttattttatac 1680  
agaattatac gcttttacta gttatagcat tcggtatctt tttctgggta actgccaaac 1740  
caccacaaat ttcaagtttc catttaactc ttcaacttca acccaaccaa atttatttgc 1800  
ttaattgtgc agaaccactc cctatatctt ctaggtgctt tcattcgttc cgaggtaaga 1860  
aaagattttt gtttctttga atgctttatg ccactcgttt aacttctgag gtttgtggat 1920  
cttttaggcg actttttttt tttttgtatg taaaatttgt ttcataaatg cttctcaaca 1980  
taaactctga caaagagaag gaattttacc aagtatttag gttcagaaat ggataatttt 2040  
cttactgtga aatatcctta tggcagggtt tactgttatt tttcagtaaa cccggg 2096

<211> 24

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (24)

<223>

<400> 101

gagctcgaat tctctgaaaa ggag

24

<210> 102

<211> 25

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (25)

<223>

<400> 102

aagctttata gagaatgtat aagat

25

<210> 103

<211> 1222

<212> DNA

<213> Lycopersicon esculentum

&lt;220&gt;

&lt;221&gt; Promotor

&lt;222&gt; (1) .. (1222)

&lt;223&gt;

&lt;400&gt; 103

gagctcgaat tctctgaaaa ggagcaccat atttgccgca ctgtggttca tatttccaag 60  
tacatttaga tgaactatat catcagattg aaagggttatt gtataatcaa tccagtggat 120  
tctcgttctg gcacctttag aagtacatgt gcggaaaaga atgataaggt ttgtattggt 180  
gttgacaaaag cctgttgctt ttctcatttg taaatgttct gaacgactcc taaattactc 240  
ttaagggtga aggtcttccg tgctgttttg taaatataat gctgtgccgt gacttacctt 300  
ttgtaccatt tgttcaaag tatggcctga acaccagggt tgtcaaaaat gtctcatgcc 360  
cgttttattg gtctgaaaat ggctgtatgc caaattctgc cgctccacag tgagcatttc 420  
gatctactgg aaattgacca acttatttta tcacttgata actaaacaaa atcctattaa 480  
ctttaatcat acattgtatt tataccgaaa aatttatgca taactcatta aattaccttt 540  
tttagcagtc aaattctaaa tcagtttcta atttatcaaa atggctttta taggggtcca 600  
ttccactaa tatacctgcc gtccatgcac tgactacaaa acaaatacct cactatgttt 660  
gttagtgctt ggtaataataa aaccttttct tttatgagaa agttcaccga gaataatttt 720  
ctatttgtgg cataatagta tatagtgcag attgacaaga atttaatttt gcagttgggc 780  
acatgaacaa ttttcctcaa agttgtagaa agtacttttc attttcttgt caccgaaaat 840  
tatttataat tgaaattaaa accgaatgag ctgcaagatt caagtcgaat tttcaaaaga 900  
attgaccaag aaaaaattca aaaatatccc ccacccccta ccaaacacat cctaaagtga 960  
ggtatagact gggactggga ttgggaaaag ggtaaaatgc tttcactagc ttagcaaaga 1020  
ttccactttg ttagctatct ttctttctca tttccttttt tctttttctt ttttttgta 1080  
tataagccaa agtaggtacc caaaagcatc aatattttgt attgcttggt gattcctctg 1140  
tagtccagta tttcattttc tacaagttcc acctccctcc ataattaacc attatcaatc 1200  
ttatacatc tctataaagc tt 1222

&lt;210&gt; 104

<211> 30

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (30)

<223>

<400> 104

gagctctaca aattaggggtt actttattca

30

<210> 105

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<221> Primer

<222> (1) .. (32)

<223>

<400> 105

aagcttagaa ttatttccaa attccgtaaa tc

32

<210> 106

<211> 1540

<212> DNA

<213> Cucumis sativus

&lt;220&gt;

&lt;221&gt; Promotor

&lt;222&gt; (1) .. (1540)

&lt;223&gt;

&lt;400&gt; 106

gagctctaca aattaggggtt actttattca ttttcatcca ttctctttat tgttaaattt 60

tgtacattta ttcaataata ttatatgttt attacaaatt ctacttttct tattcatacc 120

tattcactca agcctttacc atcttccctt tctatttcaa tactatttct acttcatttt 180

tcacgttttt aacatctttc tttatttctt gtccacttcg tttagggatg cctaagtgtcc 240

caaatttcat ctctcgtagt aacacaaaac caatgtaatg ctacttctct ctacattttt 300

aatacaaata aagtgaacaa aaatatctat aaataaacia atatataat tttgtagac 360

gctgtctcaa cccatcaatt aaaaaatttt gttatatttc tactttacct actaaatttg 420

tttctcatat ttacctttta acccccacaa aaaaaatta taaaaagaa agaaaaagc 480

taaacctat ttaaatagtt aactataaga tcttaaaatt atctcatca gtgtatagtt 540

taattgggta ttaacttata acattatata tctatgacat atactctctc ctagctattt 600

ctcacatttt ttaacttaag aaaatagtca taacatagtc taaaattcaa acatccacat 660

gctctaattt gattaacaaa aagttagaaa tattttattta aataaaaaag actaataaat 720

atataaaatg aatgttcata cgcagaccca tttagagatg agtatgcttt cacatgctga 780

gattattttc aaaactaagg ttgtagcaat attaaatcaa taaaattatt ataaataaca 840

aaattaacct gctcgtgttt gctgtatatg ggaggctaca aaataaatta aactaaagat 900

gattatgttt tagacatttt ttctatctgt attagtttat acatattaat tcaggagctg 960

cacaacccaa ttctattttc gttccttggt ggctggggtt ctcaaaagg tcaatagtca 1020

atattagggt ttattggact ttaataagta tcaaacaaat ctatgtgtga acttaaaaat 1080

tgtattaaat atttagggta acctgttgcc gtttttagaa taatgtttct tcttaataca 1140

cgaaagcgta ttgtgtattc attcatttgg cgcctcacat gcttcgggtg gctcgtttta 1200

gtctctgcct tctttgtata ttgtactccc cctcttccca tgccacgtgt tctgagctta 1260

acaagccacg ttgctgtcca ttgccaacaa agtcatttta acttcacaag gtccgatttg 1320

acctccaaaa caacgacaag tttccgaaca gtcgcgaaga tcaagggtat aatcgtcttt 1380

```
ttgaattcta tttctcttta tttaatagtc cctctcgtgt gatagttttt aaaagatttt 1440
taaaacgtag ctgctgttta agtaaatccc agtccttcag tttgtgcttt tgtgtgtttt 1500
gtttctctga tttacggaat ttggaaataa ttctaagctt 1540
```